VOL. IX 1918

THE JOURNAL

OF

Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

EDITORS:

W. C. BAGLEY,
Columbia University.

C. E. SEASHORE,
University of Iowa

J. CARLETON BELL, Managing Editor.
Brooklyn Training School for Teachers.

GUY MONTROSE WHIPPLE,
Carnegie Institute of Technology.

BALTIMORE, MD.
WARWICK & YORK, INC.
1918

TABLE OF CONTENTS

INDEX TO NAMES

Authors of original articles, communications or discussions have names printed in CAPITALS, and the page numbers in **bold face** type. Authors reviewed, contributors of reviews, and authors mentioned have page numbers in Roman.

INDEX TO SUBJECTS

VOLUME IX JANUARY, 1918 NUMBER 1

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

SOME MATHEMATICAL ASPECTS OF THE BINET-SIMON TESTS

FRANCIS N. MAXFIELD
Public School Clinic, Newark, N. J.

Since the Binet-Simon Scale for Measuring the Intelligence of Young Children was first issued, a number of translations into other languages with necessary adaptations of tests have been made and these have been followed by various revisions. These adhere in their main features and plan to the original idea of Binet and Simon.

This was to set tests for different age levels and, by observing the child's successes and failures, to give him a net score for the series which has been called the child's "Mental Age." This score has been taken as an index of the child's intelligence or intellectual status; and comparison of this score with his "Chronological Age" has been used as diagnostic of feeble-mindedness, mental normality, or mental precocity.

The Goddard revision (1911) of the Binet-Simon scale (3)[1] has been used more extensively than any other in this country and may be taken as typical. Five tests are set for each level from 3 to 12 inclusive; and standards for giving these and for scoring success or failure have been more or less definitely worked out. A child is credited with a "basal age" of the highest year for which he passes all the tests. To determine his mental age one year is added for each five tests passed from the higher age levels.

Ideally, mentally normal children of a given chronological age should make a mean mental age score which is approximately equal to their chronological age. It was soon observed by those who used the Binet-Simon Scale in its earlier form that the results obtained were not equally satisfactory for normal chil-

[1] The figures in parenthesis refer to the bibliography at the end of the article.

dren of all ages. Young children tested too high, children over 12 too low, and the best results were obtained for the intermediate ages. In the revisions of the scale which followed various efforts were made to remedy this defect. Tests for other age levels were introduced (11 year tests by Goddard (3) and tests for less than 4 years by Kuhlmann (5)). Certain tests were moved up or down in the scale to different age levels, some new tests were introduced, others discarded, and a stricter technique of giving and scoring the tests was adopted.

Goddard considered the results satisfactory if the mental-age derived was not more than one year above or below the chronological age. By this criterion he considered 78 percent of his results with 1547 children "satisfactory." But his own analysis of these cases (2) showed that these satisfactory results were by no means evenly distributed, as is shown by the following figures:

Chronological Age	4	5	6	7	8	9	10	11	12	All
Number of Cases	8	114	160	197	209	201	222	166	144	1421
Number Satisfactory	5	88	146	179	189	168	170	122	75	1142
Percent Satisfactory	63	77	91	91	90	87	76	73	56	80.4

The percentages of satisfactory cases run high for chronological ages 6 to 9 inclusive but fall off at either end of the age series. Other revisions showed similar though perhaps less striking results.

Still more striking were the results with adults and adolescents. Where these results were obtained from institutional cases they were readily accepted as showing a high frequency of mental deficiency and feeble-mindedness but when rather similar results were obtained from factory operatives and high schools some of the disadvantages of the scales became more apparent. Many who made use of the Binet tests seem to have either ignored these difficulties or to have used a lower "passing mark" in interpreting the scores for diagnostic purposes.

Thorndike (10), working over Goddard's results mentioned above and comparing them with later results of other investigators, pointed out clearly the inequalities of these scales and derived a table of "true mental-ages," in which the latter coincide closely with the Binet test mental age up to ten years. Thenceforward the true mental ages for each Binet score are increasingly advanced, giving 13.8 as the true mental age for a Binet score of 11.8. Too little attention seems to have been given to the

reason for these inequalities which lies in the very nature of the scales. A boy of 7 has, in Goddard's revision, some theoretical chance of passing 25 tests beyond the tests set for the seven year level. A boy of 11, on the other hand, has a chance of passing only 5 tests beyond those set for the eleven year level. It is obvious that, unless these 5 tests are easier for the twelve year old boy than any 5 of the 25 tests are for the seven year old boy, the seven year old boy has the better chance of securing credit for tests above his chronological age.

The Stanford Revision (8), prepared by Prof. Lewis N. Terman of Leland Stanford Jr. University and recently issued, is a serious attempt to overcome some of the chief disadvantages of the earlier revisions. In this series there are six tests set for each year level from 3 to 10 years inclusive. Success in each of these tests above the basal age is counted as two months toward the mental-age score. For 12 years there are eight tests, each counting three months, for 14 six, each counting four months, for 16 and 18 there are in each case six tests counting as five and six months respectively. The method and technique of this revision have been worked out much more carefully and stated more clearly than in any earlier revision.

In interpreting the significance of the mental age scores of children of the same or different ages or of the same child at different ages obtained from this revision the intelligence quotient or ratio of mental age score to chronological age ($\times$ 100) has been found convenient. The mean intelligence quotient of each age group tested by Terman approximates 100. But this scale also in many cases gives too high a mental age score for the very young children and too low a score for those over 12. It is the purpose of this paper to suggest one method of theoretical treatment which should be useful:

1. In comparing the results obtained with children of different ages from any revision and in interpreting the significance of these mental-age scores or intelligence quotients.

2. In comparing the validity of the results obtained by the use of different revisions.

3. In predicting the results which may be expected from still further revision or alteration.

If the individual tests of the Stanford Revision are given to normal seven year children or the so-called unselected[2] children

[2] The definition of "unselected" is uncertain, but this does not affect the argument of this paper.

who are not suspected of marked mental deficiency we find that the average relative frequency with which these children pass these tests begins with a frequency approaching unity for the four year tests and that this percentage decreases steadily for the tests set at higher age levels. If these frequencies are known, it becomes possible to compute the theoretical distribution of basal ages; *i. e.*, to find the probable number who will have a basal age of four years, five years, etc., respectively.

Let us assume for purposes of simplification that this frequency is the same for each of the tests given at any age level (six in the case of the Stanford Revision). The argument remains the same even where they vary considerably. Assume for purposes of argument that the following results are obtained (actual frequencies would have been used but Terman's results had not yet been published when this was written).

Average of relative frequency for each of	4 year tests	0.98
" " " " " " "	5 " "	0.95
" " " " " " "	6 " "	0.85
" " " " " " "	7 " "	0.75
" " " " " " "	8 " "	0.50
" " " " " " "	9 " "	0.30
" " " " " " "	10 " "	0.10
" " " " " " "	12 " "	0.05

The probability of a given seven year old child having a basal age of four years will take into account (1) the probability that he will pass all of the six tests for the fourth year $(0.98)^6$, and (2) the probability that he will not pass all the tests for any subsequent year. The probability that he will not pass all of the five year tests will be $1-(.95)^6$ and so on for the higher years. The probability P_4 that 4 will be his basal age may be computed as follows:

$$P_4 = (.98)^6\,[1-(.95)^6]\,[1-(.85)^6]\,[1-(.75)^6]\,[1-(.50)^6]\,[1-(.30)^6]\,[1-(.10)^6]\,[1-(.05)^6]$$

Similarly we have:

$$P_5 = (.95)^6\,[1-(.85)^6]\,[1-(.75)^6] \text{ — etc.}$$
$$P_6 = (.85)^6\,[1-(.75)^6]\,[1-(.50)^6] \text{ — etc.}$$

On the basis of the above table we have:

$$\begin{aligned}
P_4 &= (.886)\,(.265)\,(.623)\,(.822)\,(.984)\,(.999) &&= 0.118\\
P_5 &= (.735)\,(1623)\,(.822)\,(.984)\,(.999) &&= 0.370\\
P_6 &= (.377)\,(.822)\,(.984)\,(.999) &&= 0.305\\
P_7 &= (.178)\,(.984)\,(.999) &&= 0.175\\
P_8 &= (.016)\,(.999) &&= 0.016\\
P_9 &= (.001) &&= 0.001\\
P_{12} &= &&= 0.000\\
& &&\overline{\quad 0.985}
\end{aligned}$$

As a check on this calculation we have:

$$P_0 = (.114)\ (.265)\ (.623)\ (.822)\ (.984)\ (.999) = \frac{0.015}{1.000}$$

P_0 being the probability that no one of these basal ages will be attained.

If the average percentages for the six tests of each different age level for children of a given age be respectively p_4, p_5, p_6, p_7, etc., our calculations become:

$$P_4 = p_4{}^6\ (1-p_5{}^6)\ (1-p_6{}^6)\ldots.(1-p_k{}^6)$$

$$P_5 = p_5{}^6\ (1-p_6{}^6)\ (1-p_7{}^6)\ldots..(1-p_k{}^6) \text{ etc., and we have:}$$

$$P_n = p_n{}^6\ (1-p_{n+1}{}^6)\ (1-p_{n+2}{}^6).(1-p_{k+1}{}^6)\ (1-p_k{}^6) \quad \text{(A)}$$

If we wish to be more exact we may observe the frequency with which children of a given age pass each test and take the product of the six values for any age level and substitute this product for p^6. This will not materially change the distribution unless the mean variation is large and the argument remains the same. Suppose these products be represented by q_4, q_5, q_6, ... q_k and we then have $P_n = q_n\ (1-q_{n+1})\ (1-q_{n+2})\ .\ .\ .\ (1-q_{k-1})\ (1-q_k)$ (B).

Having determined the distribution of basal ages we may proceed to the calculation of the probabilities that tests above this basal age will be passed.

If the frequency with which each of these tests are passed is known we may calculate the probability that a certain mental age will be attained.

Assuming, for purposes of illustration, that these frequencies are the same for seven year old children for all six tests of any age level, our calculation proceeds as follows. The probability that all six of the tests will be passed may be ignored since in that case the child would be included among those whose basal age is of this or of a subsequent year. We are therefore concerned with the probability that exactly five, four, three, two, one, or none of this group will be passed.

Suppose each of the tests of any age group be passed with a frequency p, the probability

that exactly 5 will be passed	$= 6p^5\ (1-p)$,	
" " 4 " " "	$= 15p^4\ (1-p)^2$,	
" " 3 " " "	$= 20p^3(1-p)^3$,	
" " 2 " " "	$= 15p^2\ (1-p)^4$,	(C)
" " 1 " " "	$= 6p\ (1-p)^5$, and	
" " 0 " " "	$= (1-p)^6$	

As a check on these calculations we may add p^6, the probability that all six will be passed, to the sum of these probabilities and our result should be equal to unity. Suppose, for example, that 75 percent. of seven year old children pass each of the seven year tests and we have the probability

that	exactly	5	tests	will	be	passed	= 0.356
"	"	4	"	"	"	"	= 0.297
"	"	3	"	"	"	"	= 0.132
"	"	2	"	"	"	"	= 0.033
"	"	1	"	"	"	"	= 0.004
"	"	0	"	"	"	"	= 0.000
							0.822
"	"	6	"	"	"	"	= 0.178
							1.000

If the probabilities for all six tests are not the same, it will not seriously change our results to use the mean probability for the six. Or to be more exact, if these probabilities be p_1, p_2, p_3, p_4, p_5, and p_6 we have the probability that exactly five tests will be passed equal to $p_1\, p_2\, p_3\, p_4\, p_5\, (1-p_6) + p_1\, p_2\, p_3\, p_4\, p_6\, (1-p_5) + p_1\, p_2\, p_3\, p_5\, p_6\, (1-p_4) + p_1\, p_2\, p_4\, p_5\, p_6\, (1-p_3) + p_1\, p_3\, p_4\, p_5\, p_6\, (1-p_2) + p_2\, p_3\, p_4\, p_5\, p_6\, (1-p_1)$. (D) Similarly the probability that exactly four tests will be passed is equal to:

$$p_1p_2p_3p_4\,(1-p_5)\,(1-p_6) + p_1p_2p_3p_5\,(1-p_4)\,(1-p_6) + p_1p_2p_3p_6\,(1-p_4)\,(1-p_5) +$$
$$p_1p_2p_4p_5\,(1-p_3)\,(1-p_6) + p_1p_2p_4p_6\,(1-p_3)\,(1-p_5) + p_1p_2p_5p_6\,(1-p_3)\,(1-p_4) +$$
$$p_1p_3p_4p_5\,(1-p_2)\,(1-p_6) + p_1p_3p_4p_6\,(1-p_2)\,(1-p_5) + p_1p_3p_5p_6\,(1-p_2)\,(1-p_4) +$$
$$p_1p_4p_5p_6\,(1-p_2)\,(1-p_3) + p_2p_3p_4p_5\,(1-p_1)\,(1-p_6) + p_2p_3p_4p_6\,(1-p_1)\,(1-p_5) +$$
$$p_2p_3p_5p_6\,(1-p_1)\,(1-p_4) + p_2p_4p_5p_6\,(1-p_1)\,(1-p_3) + p_3p_4p_5p_6\,(1-p_1)\,(1-p_2) \quad \text{(E)}$$

And we have similar series for three, two, one, and no tests, respectively.

Suppose that $p_1 = .7$, $p_2 = .7$, $p_3 = .75$, $p_4 = .75$, $p_5 = .8$ and $p_6 = .8$ then the probability

that	exactly	5	tests	will	be	passed	= 0.357
"	"	4	"	"	"	"	= 0.298
"	"	3	"	"	"	"	= 0.132
"	"	2	"	"	"	"	= 0.032
"	"	1	"	"	"	"	= 0.004
"	"	0	"	"	"	"	= 0.000
							0.823

$p_1p_2p_3p_4p_5p_6$ (the probability that all six will be passed) = 0.176

0.999

It is apparent that these results do not differ materially from those derived from assuming these probabilities all equal to the average, *i. e.* 0.75, as above.

If, therefore, one is to determine the expectations of mental age scores for the seven year children of each basal age group, he must calculate the probabilities that a certain total number of tests beyond the basal year will be passed. To do so, he will calculate as above the probabilities that 5, 4, 3, etc., tests respectively will be passed from each age group. He will tabulate these results and proceed to calculate the probabilities that a certain number will be passed from all groups above the basal age, beginning at the highest.

If the probabilities that 5, 4, 3, etc., tests are passed from the ten year group be respectively q_5, q_4, q_3, etc., and from the nine year group be r_5, r_4, r_3, etc., and the probabilities that 10, 9, 8, etc., respectively be passed from the two groups be x_{10}, x_9, x_8 etc., we have:

$$x_{10} = q_5r_5$$
$$x_9 = q_5r_4 + q_4r_5$$
$$x_8 = q_5r_3 + q_4r_4 + q_3r_5, \text{ etc., and}$$
$$x_1 = q_1r_0 + q_0r_1$$
$$x_0 = q_0r_0$$

As a check on these calculations, we have $\Sigma_q \, \Sigma_r$ etc., $= \Sigma_x$ where Σ_q is the sum of $q_5 + q_4 + \ldots q_0$, etc. In a similar way we may calculate the probabilities that 15, 14, 13, etc., tests respectively will be passed from the three year groups 8, 9, and 10. If these probabilities be y_{15}, y_{14}, y_{13}, etc., and those for the 8 year group be s_5, s_4, s_3, etc., we have:

$$y_{15} = x_{10}s_5$$
$$y_{14} = x_9s_5 + x_{10}s_4$$
$$y_{13} = x_8s_5 + x_9s_4 + x_{10}s_3, \text{ etc.,}$$
and
$$y_1 = x_0s_1 + x_1s_0$$
$$y_0 = x_0s_0$$

Again we have as a check:

$$\Sigma_x \, \Sigma_s = \Sigma_y$$

On the basis of the distribution of basal ages derived in our first table, we may proceed to determine the relative distribution of each group according to mental age. Since none attains a basal age of 10, we may assume that those represented by P_0, *i. e.*,

those not attaining the basal ages given, have a basal age of three years. If we distribute each basal age group according to the relative expectation of passing additional tests we get the following result:

Distribution of 7 year children according to mental-age score (Stanford Revision)

Mental age score yrs.	mo.	Basal ages 3	4	5	6	7	8	9	Total
9	0							.001	.001
	10						.001		.001
	8					.001	.002		.003
	6				.001	.005	.004		.010
	4			.001	.005	.013	.005		.024
	2			.004	.013	.025	.003		.045
8	0		.001	.012	.028	.036	.001		.078
	10		.004	.027	.047	.038			.116
	8		.008	.047	.059	.030			.144
	6	.001	.014	.065	.059	.018			.157
	4	.002	.020	.070	.046	.007			.145
	2	.003	.022	.061	.028	.002			.116
7	0	.003	.020	.043	.013	.000			.079
	10	.003	.014	.024	.005				.046
	8	.002	.009	.011	.001				.023
	6	.001	.004	.004	.000				.009
	4	.000	.002	.001					.003
	2	.000	.000	.000					.000
6	0								
		.015	.118	.370	.305	.175	.016	.001	1.000

It will be observed that this computation is effected in the case of the basal age 7, for example, by multiplying each of the expectations for passing tests from the 8, 9, and 10 year levels (y_{15}, y_{14}, y_{13}, etc.) by P_7 divided by the sum of these expectations, *i. e.*, $P_7 \div \Sigma_y$. Thus $\frac{P_7}{\Sigma_y} y_5 = 0.038$, the expectation that 7 year children having a basal age of 7 will pass five tests from the higher age levels and therefore attain a mental-age score of 7 yrs., 10 mos. This gives a "normal" distribution of mental-age scores with a mode of seven years and six months. This would suggest that the assumed frequencies are too high but this does not affect the argument.

In the case of the Goddard revision the determination of the theoretical distribution of mental-age scores would be made in much the same way. Using a similar notation and argument we have formulae, analagous to (A) and (B) above, for the distribution of basal ages as follows:

$$P_n = p^5_n\,(1-p^5_{n+1})\,(1-p^5_{n+2})\ldots\ldots\ldots\ldots\ldots(1-p^5_k) \text{ or}$$
$$P_n = q_n\,(1-q_{n+1})\,(1-q_{n+2})\ldots\ldots\ldots\ldots\ldots\ldots(1-q_k)$$

Here the probability that five tests will be passed from any age level above the basal age is excluded and for the probability that exactly four tests will be passed from any age level we have (where the relative frequencies for passing each of the tests of this age level are all equal to p) $5p^4 (1-p)$,
for exactly 3 tests $10p^3 (1-p)^2$,
" " 2 " $10p^2 (1-p)^3$,
" " 1 " $5p (1-p)^4$, and
" " 0 " $(1-p)^5$.
These are, of course, analagous to (C) above.

Where these probabilities for any age level are not equal (p_1, p_2, p_3, p_4, and p_5) we have the probability that exactly four tests will be passed equal to $p_1p_2p_3p_4 (1-p_5) + p_1p_2p_3p_5 (1-p_4) + p_1p_2p_4p_5 (1-p_3) + p_1p_3p_4p_5 (1-p_2) + p_2p_3p_4p_5 (1-p_1)$, analagous to (D) above. Similarly the probability that exactly three tests will be passed is equal to $p_1p_2p_3 (1-p_4) (1-p_5) + p_1p_2p_4 (1-p_3) (1-p_5) + p_1p_3p_4 (1-p_2) (1-p_5) + p_2p_3p_4 (1-p_1) (1-p_5) + p_1p_2p_5 (1-p_3) (1-p_4) + p_1p_3p_5 (1-p_2) (1-p_4) + p_2p_3p_5 (1-p_1) (1-p_4) + p_1p_4p_5 (1-p_2) (1-p_3) + p_2p_4p_5 (1-p_1) (1-p_3) + p_3p_4p_5 (1-p_1) (1-p_2)$. This is analagous to (E) above and we have similar series for two, one, and no tests.

For illustration of this method with the Goddard revision I am indebted to Mr. E. A. Doll of the Vineland Laboratory for empirical frequencies of so-called unselected children on each of the tests of this revision. I have made use of the data for the seven year group (39 children with average chronological age 7.4 years), the ten year group (68 children with average chronological age 10.5 years) and the twelve year group (69 children with average chronological age 12.5 years).

A theoretical distribution of mental-age scores by this method based on the observed mean frequencies with which children of these groups passed the individual tests gives results which correspond closely with the actual results in the case of these children, as will be observed from the following comparative table:

Age Groups	Number of Cases	Mean Chronological Age	Mean Mental-Age Scores	
			Actual	Calculated
7	39	7.4	7.2	7.3
10	68	10.5	10.6	10.7
12	69	12.5	10.8	10.9

It has been customary in working upon any revision or modification of the Binet-Simon tests to rescore individual records or re-test children according to the revised or modified scale. It would appear that, where no new tests were introduced, as for example in Doll's Abbreviated Scale, it would be a fairly reliable method to test the new scale by distributing the mental-age scores theoretically on the basis of frequencies already observed. Where new tests were introduced, it would be only necessary to determine empirical frequencies for these tests and introduce them into a theoretical treatment along with the data previously obtained for the tests which were not new.

As an example of this treatment Doll's Abbreviated Scale (1) gives for the 39 children of the seven year group mentioned above a mean mental-age score of 7.0 and for the 68 children of the ten year group a mean mental-age score of 9.6. Doll's actual results on other children give 7.0 and 9.7 for these age groups.

In the case of the Stanford Revision this method must be somewhat modified where tests above the 10 year level are involved. The argument, however, remains the same. In calculating the distribution of basal ages one must use p^8 and $(1—p^8)$ in place of p^6 and $(1—p^6)$ for the 12 year level. In calculating the expectation of passing beyond the 10 year level for distributing the mental age scores a new complication arises. The tests are weighted differently since a different value in months is given to the tests of different age levels. For this reason it is convenient to reduce each to months. A boy passing 5 eighteen year tests and 4 sixteen year tests scores (5×6) + (4×5) or 50 months. This reduction to months may be continued into the age levels below 12 or after deriving the expectation from the 12, 14, 16, and 18 year levels one may reduce them to the equivalent of two month tests. Thus for 80 and 81 one may substitute 40, and for 78 and 79, 39, since the final expectations and distributions derived are to be on a two months' basis. Of course p_{40} in this case would equal $p'_{80} + p'_{81}$, and $p_{39} = p'_{78} + p'_{79}$.

The method can also be easily adapted to the abbreviated Stanford Scale (tests marked by asterisk) by an appropriate weighting of the tests involved. The theoretical correlation between this and the complete scale can thus be easily determ-

ined on the basis of empirical frequencies without the necessity of rescoring individual results or of testing additional children.

The term "mental-age score" rather than "mental-age" has been used in this treatment because it is objective and concrete, a qualitative and quantitative measure of a child's reactions to a series of tests given and scored in a standardized manner. Confusion has arisen because the term mental-age has been used by many writers in two senses which are not easily distinguished by the reader. On the one hand it is held to signify a quantitative expression of the child's intellectual status or intelligence level in terms of the average child. To say in this sense that a seven year child has a mental-age of eight years is to say that the child's intellectual status is equivalent to that of an average eight year child and above that of the average child of his own age. On the other hand the term mental-age has been used to signify the score attained by the child in a series of Binet-Simon tests. Even if one grants the psychological advantage of the concept mental-age in the first of these meanings, which may well be challenged except for popular use, one is only justified in using these terms interchangeably by assuming a perfect correlation between this score as obtained in the case of any child and the measure of that complex more-or-less undefined something which we call his intelligence or intellectual status. It is much more in keeping with our present limited knowledge of intelligence not to make this assumption and, except for popular use, to avoid the use of the term mental-age. As Witmer has pointed out (11), these tests show a child's "performance level" and one should speak of his mental age score rather than his mental age. It is confusing to use the term mental age in any sense other than that of a score or performance level in making the statement that "normal five year children have a mental age of six years." (Approximately the average score by the Goddard revision). Yerkes has overcome this ambiguity by scoring the child's performance on his point scale. Terman's term "intelligence quotient" also will avoid confusion if one does not press too strongly the assumption that the average child's intelligence quotient should be 100, or the idea that intelligence alone has been tested.

In other words the result of the child's performance is a score. To judge of the child's intellectual status or intelligence it is

necessary to interpret this objective result. His high score may be due to good health, good training, good intelligence in one or more performances involving different mental abilities or functions or in a combination of all of these. A low score must also be interpreted rather than accepted without question as an infallible index of low intelligence. It may, for example, be due to the arrangement of the tests in the scale as when a twelve year old boy is tested by the Goddard revision without using tests beyond the 12 year level. The chances are against his obtaining an I. Q. of 100 as he has no tests offered him for raising his score above 12 years.

Otis has already pointed out (6 and 7) that the relative frequencies obtained by those who have administered Binet-Simon tests are objective data which may be treated mathematically. These methods, whether applied to the Goddard, Stanford, or similar revisions, or adapted to the Yerkes-Bridges scale, should be developed so that these data may be considered on their own merits and not on any *a priori* assumption of "the validity of the scale," the nature of mental deficiency, etc.

REFERENCES

1. Doll, Edgar A. *A Brief Binet-Simon Scale.* The Psychol. Clinic, 1917, **11**, 197–211; 254–261.
2. Goddard, Henry H. *Two Thousand Normal Children Measured by the Binet-Simon Measuring Scale of Intelligence. Ped. Sem.*, 1911, **18**, 232–259.
3. Goddard, Henry H. *Articles on Goddard 1911 Revision of Binet-Simon Scale.* Training School Bull., 1911, **8**, 52–62; 1913, **10**, 23–30; 1913, **10**, 44–45.
4. Kohs, Samuel C. *The Binet-Simon Measuring Scale for Intelligence: An Annotated Bibliography.* J. of Educ. Psychol., 1914, **5**, 215–224, 279–290, and 335–346.
5. Kuhlmann, F. *A Revision of the Binet-Simon System for Measuring the Intelligence of Children.* J. of Psycho-Asthen., No. 1, Monogr. Suppl., 1912, pp. 41.
6. Otis, Arthur S. *Some Logical Aspects of the Binet Scale.* Psychol. Rev., 1916, **23**, 152, and 165–179.
7. Otis, Arthur S. *A Criticism of the Yerkes-Bridges Point Scale, with Alternative Suggestions.* Journ. of Edu. Psych., 1917, **8**, 129–150.
8. Terman, Lewis M. *The Measurement of Intelligence.* Boston: Houghton Mifflin Co., 1916, pp. xviii+362.
9. Terman, Lewis M. *et al. The Stanford Revision and Extension of the Binet-Simon Measuring Scale of Intelligence.* Baltimore: Warwick and York, 1917.
10. Thorndike, Edward L. *The Significance of the Binet Mental Ages.* The Psychol. Clinic, 1914, **8**, 185–189.
11. Witmer, Lightner. *On the Relation of Intelligence to Efficiency.* The Psychol. Clinic, 1915, **9**, 61–86.

THE MEASUREMENT OF INTELLIGENCE: SIX HUNDRED AND FIFTY-THREE CHILDREN EXAMINED BY THE BINET AND PORTEUS TESTS

S. D. PORTEUS

Superintendent of Special Schools, Fitzroy, Melbourne, Australia

Since the publication of my tests in 1915[1] a considerable amount of work has been done with them, chiefly in the direction of making further comparisons of the results obtained by their use with those got by the use of the Binet-Simon scale. They have been applied to several groups of normal children as well as to feeble-minded. It is the object of this paper to report these results and to offer some observations that have arisen through the extension of their use.

Re-Adjustment of Scoring

When the tests were first arranged they were graded, somewhat loosely, for the ages from three to thirteen years. This grading was for convenience only. It was not claimed that the age allotted to the highest test which was passed represented exactly the child's mental age. All that the tests were intended to show was that, in certain capacities, notably in prudence, forethought, ability to sustain attention, in power to profit by experience as shown by improvement through the practice afforded in the tests themselves, the child had reached a certain stage of development. It did not follow that this level corresponded with that of the child's general intelligence, though, in the majority of cases (between sixty and seventy per cent.) it was shown that the development of these capacities did proceed *pari passu* with that of general intelligence.

When the tests are used as an aid to diagnosis, the fact that the child quickly improves with practice is important. Hence, when the subject of the tests was suspected to be mentally deficient, it did not matter greatly whether the child barely succeeded or passed the tests with comparative ease. Yet this

[1] Journal of Psycho-Asthenics, June, 1915.

scraping through the tests by taking the full number of trials allowed, whilst it argued the possession of the power to improve rapidly with practice, certainly showed that the child lacked mental alertness. In order that the child's capacities might be exhibited in the fairest light, the tests for the lower years had been made a little easy for the mental ages allotted to each. In order, however, to make it unlikely that morons should pass the highest tests, these latter were made relatively more difficult. Hence, when diagnosis as to mental deficiency was not called for—as in the case of normal children—some adjustment was necessary. This was in order to make a distinction between the mentally alert and the slower, duller child who succeeded in passing the tests, but with frequent errors.

This adjustment, then, consists in counting the number of tests which the child passes on the last trial allowed by the test conditions: (two trials are allowed in tests up to, and including the eleven-year test, three trials in the twelve, and in the thirteen-year tests). If two tests are passed on the last trial, one year is deducted; if three tests, one year and a half; if four tests, two years are deducted from the age allotted to the highest test the child passes. The remainder is the mental age. Should a child fail outright in a test but succeed in a higher one, a year is deducted. Thus, for instance, a child passes in the 9-year test (two trials), fails in the 10-year test (two trials), passes in the 11-year test (one trial), succeeds in the 12-year test (three trials), and finally fails in the 13-year test. Twelve is the highest test passed. A year is deducted for the failure in the 10-year test, and another year for the two tests (nine and twelve) in which the child passed on the last trial. The mental age is twelve minus two years.

This adjustment evidently increases the difficulty of the series by making the scoring conditions harder. It also has the effect of minimizing some inequalities in the steps of the series. For instance, if a child fails in the 10-year test but succeeds in the 11-year, the deduction of a year is equivalent to transposing the positions of the two tests in the series.

Though individual tests, such as those for 10 years and 12 years, are somewhat difficult, it has not been considered advisable to alter them until sufficient data have been collected to make a final revision possible. In the meantime, the foregoing

adjustment makes the series more serviceable. By altering the scoring rather than the tests themselves, results previously obtained are still available for comparison.

Careful observation by the examiner during the application of the test is very necessary. Many interesting individual differences in the methods of attacking the problems may then be noticed. The child of impulsive and headstrong temperament does not attempt any preliminary survey of the design, but attempts to rush through the maze, usually falling into the first trap that presents itself. Nor does the failure deter it from making similar errors. Such a child will very often see its error immediately it is committed, but it seems quite impossible for it to pass a wrong turning without committing itself to the wrong course. The repetition of this kind of action seems almost to savor of obstinacy. The inference is that such children habitually proceed to action and consider the effect afterwards. They leap before they look.

Another child betrays irresolution and indecision of character by its action. It seems to have great difficulty in deciding which course to take. It studies the test too much. The possible mistakes obtrude themselves upon consciousness to such a degree that it cannot hold the clear course in view, so that it ends by confusing itself and falls into the very error it is so anxious to avoid. To such children the 11-year test seems to offer peculiar difficulty, whereas others find it particularly easy.

Over-confidence causes some to fall into error. The over-confident child may be distinguished from the child who is habitually impulsive and impetuous. The former learns prudence by its first error, the latter does not. It will be observed that allowance is made for over-confidence, as the first mistake is not penalized. The most intelligent attitude towards the test is that which is characterized by foresight in a preliminary planning of the course, together with a prudent following out of the plan. In some, impetuosity of action nullifies the good effect of preliminary study. Such children appear to trust too much to memorizing the whole course to be pursued, instead of using careful observation whilst at work.

The most common source of error is heedlessness, carelessness, inability to allow thought to precede action. A child who possesses this fault may quite possibly appear brilliant as regards

educational attainment. School examinations and tests of attainment such as the Binet show these children in a too favorable light. We are prone to over-rate the glib-tongued, quick-witted person who may possess, at the same time, an infirmity of temperament so grave as to render success in after life unlikely. Irresolution, impetuosity, heedlessness, lack of staying power, count for comparatively little in the childish period. It is only when later life brings real problems to the individual, that these faults of disposition, previously overlooked, become apparent. Hence it is that fortunate prognostications as to the success of some of our brilliant pupils are unfulfilled.

Performance of the Feeble-minded

In the case of the feeble-minded, school progress weighs too heavily with us when we attempt diagnosis. There is one well-marked type of child in whom educational aptitudes are at least moderate, but there is such a feebleness of will energy, such a morbid self-consciousness and fearfulness of failure, that he is quite unfitted to compete with his fellows, but remains in a state of dependency throughout life. Surely, in the evaluation of intelligence, such traits of character are at least equal in importance to reading, spelling, and computing skill. In childhood such cases are well designated "potentially feeble-minded." Hence the importance of tests that enable us to prove the possession of the positive virtues of care and prudence, deliberative action and forethought.

Somewhat contrary to my first expectation, the feeble-minded, as compared with normals, do not show any specific differences in their methods of attacking the tests. That is to say, the young normal child makes the same mistakes and in the same manner as do the mentally deficient who are older. Some undoubtedly feeble-minded children exhibit care and prudence, though the majority are impulsive and unable to profit by experience. But, as before mentioned, it is this very heedlessness—though shown in a lesser degree—which is the most fruitful source of failure with children of ordinary development. There is, then, no particular incapacity which may be considered peculiar to the feeble-minded. The correlation between the results obtained from the use of the Binet and of my own tests for a group of 200 mentally deficient is very little higher than that

for a corresponding group of normal school children. The view previously taken, that the tests are peculiarly adapted for the mentally deficient holds good only from the standpoint that the proof of the possession of the capacities tested by my series is so important both in diagnosis and prognosis. The exhibiting of the feeble-minded child's one talent is relatively more important than its display in a normal child who possesses also the other nine. Apart from this consideration, the theory that incapacity for prudence, foresight, etc., is one of degree and not of kind is important from the standpoint of the psychology of feeble-mindedness. If substantiated, it would tend to prove that "feeble-minded" is not a generic term, but that the mental defective is simply a very dull normal. If there is no dividing line between mental deficiency and mere mental dullness, the search for absolutely diagnostic tests is vain. Definition must proceed through social reference to the "ability to manage oneself and one's affairs with ordinary prudence." Like the insane, the feeble-minded will be defined by what he does. Hence the importance of tests of performance, provided they give any reliable indication of habitual behavior and disposition. In this connection, the application of my own and similar tests to social delinquents should give interesting results.

Comparison of Binet and Porteus Tests

For purposes of comparison, the Binet and my own tests were applied to three groups of children. The first consisted of two hundred feeble-minded children. A summary of the variations between the results of the two tests is shown in Table 1. The individual records are interesting but would occupy too much space to offer in detail. Table 2 gives a similar summary for a group of 190 normal children, taken in a chance distribution through the lower and middle grades of a city school. For both the above-mentioned groups Goddard's revision of the Binet scale was used. This revision was found to be eminently serviceable, except for the upper years. The test as arranged for 10 years appeared too difficult for Australian children of that age. If only the common coins were used in the recognition test, the question was too easy, since nearly all 9-year children knew them. If the more uncommon gold coins—uncommon, that is, from the child's experience—were used,

very few children of 10 years knew them. The two memory tests in this year are difficult, whilst several of the questions in the comprehension-of-difficult-questions test (X 4) are needlessly incomprehensible. Again, in the 12-year test, the "memory for sentences" test is certainly too hard, whilst failure in the "resistance to suggestion" test does not, I believe, signify much regarding intelligence.

Considering these imperfections in the scale, one had reason to doubt the reliability of estimates of mental age for children of the higher ages. In order to collect data for comparison for the ages 11 to 14, Terman's revision was used with a third group of children, of whom about 80 belonged to the upper grades. I am inclined to the belief that the Terman revision is more correct on the whole. Certainly, the higher age tests appear more reliable. At the time of the investigation the last-named revision had not been published, but was made available for use through the courtesy of Dr. Terman. Though the position of certain tests in the lower years is surprising, their difficulty is lessened by making the scoring conditions easier; as, for example, in certain rote memory tests, where one out of three trials suffices for a pass. The summary for this third group is given in Table 3.

Table 4 shows the results for all cases, 653 in number. The great majority of the children were tested by the writer personally, and all under his immediate supervision:

Summary of Variations

Binet and Porteus Tests

TABLE 1

Two Hundred Mentally Deficient

Variation	Number of Cases	(Goddard's Revision) Percentage
No difference	45	22.5
Half Year	48	24
One Year	61=154	30.5=77
1½ Years	18	9
2 Years	14=32	7 =16
2½ Years	7 14	3½
3 Years	6	3
3½ Years	1=14	½= 7
Total Cases	200	

Correlation: r=.7 P.e=.024.

TABLE 2

One Hundred and Ninety Normal Children

Variation	Number of Cases	(Goddard's Revision) Percentage
Half Year	45	23.6
One Year	46=137	24.2=72
1½ Years	22	11.6
2 Years	13=35	6.8=18.4
2½ Years	7	3.7
3 Years	5	2.6
3½ Years	5	2.6
4 Years	1=18	.5=9.5
Total Cases	190	

Correlation: r= .69

As will be seen by comparing Tables 1 and 2, in both groups over 70% showed variations of a year or less, whilst over 90% showed a deviation of two years or less.

The results by the two series of tests were in slightly closer agreement with the feeble-minded group than with the normal.

TABLE 3

Two Hundred and Sixty-three Cases (Terman's Revision)

Variation	Number of Cases	Percentage
No difference	49	
Half Year	82	
One Year	60=191	72.6
1½ Years	44	
2 Years	15=59	22.4
2½ Years	11	
3 Years	1	
3½ Years	1=13	5

Correlation: r=.77

TABLE 4

Six Hundred and Fifty-three Cases (All Tables)

Variation	Number of Cases	Percentage
No difference	140	
Half Year	175	
One Year	167=482	73.8
1½ Years	84	
2 Years	42=126	19.2
2½ Years	25	
3 Years	12	
3½ Years	7	
4 Years	1=45	7

In Tables 1 and 2, in reckoning the Binet age one-fifth of a year has been disregarded, two-fifths and three-fifths reckoned as half a year and four-fifths as a whole year. In this I have followed Mr. Cunningham's suggestion. For the Terman revision (Table III) up to three months has been disregarded, from four to nine months reckoned a half-year, and ten and eleven months an additional year.

From examination of the tables it will be seen that the relation between the two series is fairly constant. We may safely make the generalization that in any large group of unselected children about 70 per cent. of the total number will test by my series within one year of their mental age according to Binet. In the present investigation this has held good for each hundred tested. This fact is important, as it has a bearing on the summary dealing with sex-differences.

Assuming the general reliability of the Binet this agreement justifies the contention that, in the remaining 30 per cent. of cases, the verdict of the Porteus tests is worth considering when we are attempting the evaluation of intelligence.

Sex Differences

The following summary is based on the performances of boys and girls in the normal groups only. It gives a somewhat remarkable view of the sex differences. If the number (453) tested is sufficient to found such an important generalization upon, it shows that the performance of the girls differs distinctly from that of the boys. In the case of the former there is a closer agreement between the estimates of intelligence by the two series. Seventy-nine per cent. of the girls test within one year of their mental age according to Binet as against 67 per cent. of the boys. The limits of variability in the case of the girls are narrower; no girl tests more than 2½ years above or below her mental age. Thirteen boys show a greater variation than 2½ years.

That the girls generally speaking test lower than the boys is shown by the fact that 40 per cent. test below their Binet age, 33 per cent. above. Of the total number of boys, approximately 30 per cent. test below their Binet age, 53 per cent. above.

The conclusion that these figures appear to warrant is that in the capacities tested by my series girls do not show, on the

whole, a development equal to that of the boys. Their results tend to group closer around the Binet ages, with at least as many below as above. Boys tend to test higher, with a greater mean variation.

The number of girls (200) that was tested is certainly not so great as the number of boys (253). Yet the group results shown in the previous tables justify the expectation that a larger number would give similar results. If further investigations confirm the conclusion, an interesting sex difference will have been displayed. A possible explanation is that the puzzle element in the tests appeals more to the boys' interest than to the girls', although lack of interest is rarely apparent in either boys or girls. Probably the impulsive temperament is more often found in females. Foresight and the ability to plan are usually considered masculine attributes. It would be very interesting indeed if experimental evidence were afforded that, even in childhood, behavior reflects these sex distinctions. The figures are appended.

TABLE 5

Sex and Performance

Variation	Number of Cases	Boys	Percentage	Number of Cases	Girls	Percentage
No difference...	41			53		
Half-Year......	70			58		
One Year......	59			47		
	—			—		
		170	67.1		158	79
1½ Years......	40			26		
2 Years......	19			9		
	—			—		
		59	23.3		35	17.5
2½ Years......	11			7		
3 Years......	6			0		
3½ Years......	6			0		
4 Years......	1			0		
	—			—		
		24	9.5		7	3.5
Total Cases.. (Boys)	253		Total Cases.. (Girls)	200		

	Boys	Girls
Number Testing Above their Binet Age........................	134 (53%)	66 (33%)
Number Testing Below their Binet Age........................	78 (30%)	81 (40%)

TABLE 6

Relation to Chronological Age

1. Number who tested *above* their actual age by both series (advanced)	178
2. Number who tested *below* their actual age by both series (retarded)	129
Total	316
3. Number who tested above age by one series and below by the other	137
Grand Total	453

Results of Re-Examinations

Table 7 gives the results of another investigation showing the effects of re-examining children after a year's interval. In 1915 Mr. K. S. Cunningham examined 100 normal children by both series.[2] In 1916 Miss Rolls re-examined 47 of these children, with a view to determining whether my test could be re-applied with advantage. Results would have been more reliable if the same observer had carried out both investigations. As each, however, had worked under my immediate supervision the reliability co-efficient was probably high.

By reference to columns V and VIII of the table the year's gains or losses for each child in each series can be compared. Some children evidently profited by their previous experience in my tests, and improved their positions by from two to four years. The first 20 cases listed in the table made, however, less than a year's progress. That the improvement of the others was not wholly due to previous experience is shown by averaging the Binet gains. The first 20 gained among them 9½ years and lost 1½ years. The average net gain for these children may be reckoned as two-fifths of a year. The remaining 27 children gained a little over one year according to the Binet and two years according to my series. Some of the gains in my tests were in the direction of making up a deficiency as compared with the Binet, and in 11 cases brought the two estimates of mental age closer for 1916 than they were in 1915.

A very significant fact is that, in the case of the Binet, 10 children only out of the whole number gained one year in mental age during the 12 months, whilst nine children made a similar gain by my tests. Nine children by the Binet, and 12 by my tests, did not gain at all. This goes to show that mental development, chronologically considered, is by no means regular.

[2] K. S. Cunningham. *Binet and Porteus Tests Compared. Examination of One Hundred School Children.* Journal of Educational Psychology, 7: 1916, 552–556.

TABLE 7

Re-examination After a Year's Interval

Case No.	Age 1915	Binet 1915	Binet 1916	Binet Gain or loss	Porteus 1915	Porteus 1916	Porteus Gain or loss	Variations between tests 1915	Variations between tests 1916
1	8	10	10½	+ ½	11	10	—1	1	½
2	8½	10½	11	+ ½	12	11	—1	1½	0
3	7½	9	9	0	10	9	—1	1	0
4	6½	7	6½	— ½	7	6	—1	0	½
5	8	7	8	+1	10½	10	— ½	3½	2
6	6½	8½	10	+1½	7	7	0	1½	3
7	9	10	10	0	13	10	0	3	3
8	9	10	10½	+ ½	12	12	0	2	1½
9	10	10½	10½	0	8	8	0	2½	2½
10	5½	7	7	0	7	7	0	0	0
11	7	9	10½	+1½	10	10	0	1	½
12	8	9	10	+1	9	9	0	0	1
13	6	8	8½	+ ½	10	10	0	2	1½
14	9½	11	11½	+ ½	11	11	0	0	½
15	9½	10½	11	+ ½	11	11	0	½	0
16	12	10	10	0	10	10	0	0	0
17	6	7	6½	— ½	7	7	0	0	½
18	5	6	7	+1	6	6½	+ ½	0	½
19	6½	9	9½	+ ½	8½	9	+ ½	½	½
20	11	11	10½	— ½	10½	11	+ ½	½	½
21	11	9	11	+2	9	10	+1	2	1
22	10	8	9	+1	8	9	+1	0	0
23	6	7	7½	+ ½	5	6	+1	2	1½
24	5½	6½	7½	+1	7	8	+1	½	½
25	9½	11	11	0	12	13	+1	1	2
26	6	9½	9½	0	9	10	+1	½	½
27	7	6½	8	+1½	7	8	+1	½	0
28	6½	7½	8	+ ½	8	9	+1	½	+1
29	5½	7	7	0	6	7	+1	1	0
30	7½	9½	10½	+1	9	11	+2	½	+ ½
31	8½	9½	11	+1½	10	12	+2	½	+1
32	9	7½	7½	0	7	9	+2	½	1½
33	6½	8½	10	+1½	7	9	+2	1½	1
34	9½	9	11	+2	9	11	+2	0	0
35	9	7	8½	+1½	8	10	+2	1	1½
36	5½	6½	7½	+1	5	7	+2	1½	½
37	6½	6½	9	+2½	7	9	+2	½	0
38	7	9	10	+1	8	10	+2	1	0
39	9	8	9½	+1½	8	10½	+2½	0	1
40	7½	10½	11	+ ½	8	11	+3	—1½	0
41	7	9	9½	+ ½	8	11	+3	1	+1½
42	7½	6	7½	+1½	6	9	+3	0	1½
43	6½	9	10	+1	8	11	+3	—1	+1
44	7	8	10	+2	6	9	+3	—2	—1
45	7½	7½	8	+ ½	6	9	+3	—1½	+1
46	5	6½	8	+1½	6½	10	+3½	0	2
47	7½	9	10	+1	7½	11½	+4	—1½	+1½

The number tested is, of course, too small to found generalizations upon, but the table is certainly somewhat interesting. It is to be regretted that a greater number of Mr. Cunningham's cases were not available for re-examination. The results of this investigation are, however, sufficient to prove that my test can be re-applied with advantage, provided that a long enough interval of time is allowed to elapse between the two examinations.

Age Distribution

In the accompanying graph the actual distribution according to their mental ages of all children is shown, the Binet ages by a black line, the Porteus by a dotted one. Half-years have been disregarded, and one curve has been superimposed upon the other. As will be seen the curve for the Porteus series has a distinct "skew" for the 13-year numbers. This may be at least partially explained by the fact that this age marks the upper limit of the scale, and that some children would undoubtedly have gone beyond this limit had there been higher tests. Otherwise the distribution curve may be regarded as fairly regular. It is somewhat steeper in outline than that for the Binet ages. The actual age distribution is shown beneath. It will be seen that there was, actually, an excess of 13-year-old children.

Where Errors Occur

The maze designs that are appended are to show an analysis of the errors that were made by one hundred seventh and eighth grade children. The numerals show the number of errors that were made at each wrong turning or opening.

The large number of errors shown in the ten-year test is partially accounted for by the fact that this was the first test given throughout this investigation. Consequently, some children made mistakes through under-estimating its difficulty or through lack of practice afforded by the usual course of working through the lower tests. The proper procedure is to start with the easier tests. It is not advisable to start any children with a test higher than that for eight years.

By following the plan of marking where errors are made, data as to the relative difficulty of the individual tests can be obtained. These data could then be used in any attempt at

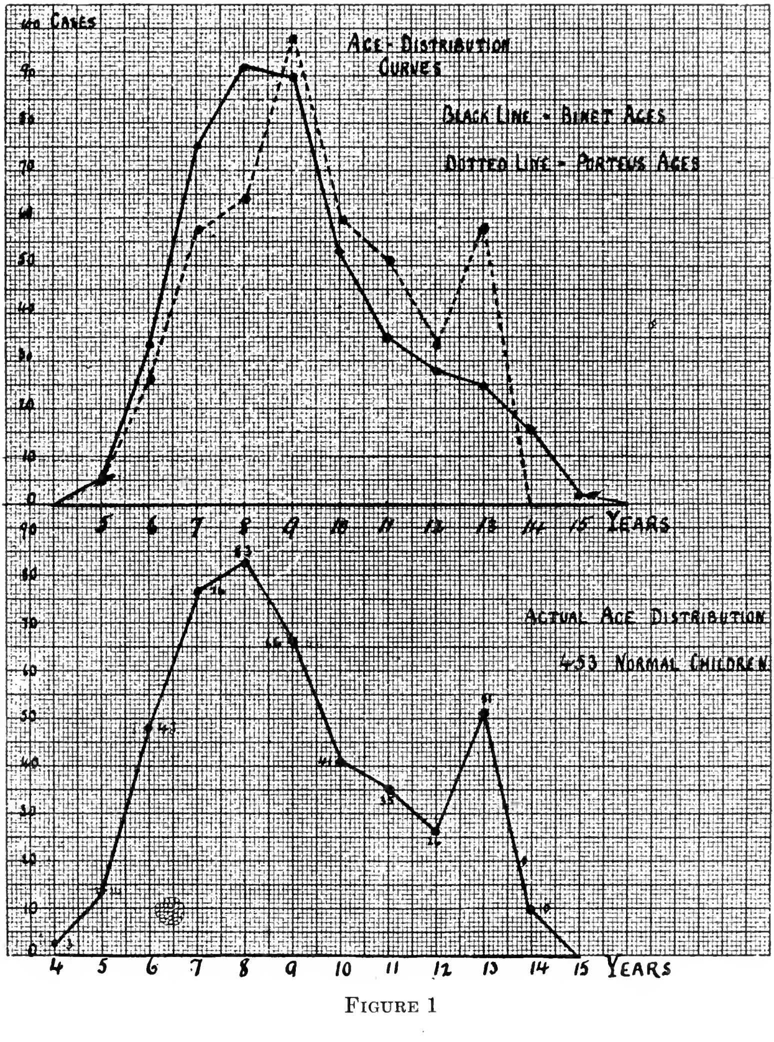

Cases
Age - Distribution
Curves
Black Line = Binet Ages
Dotted Line = Porteus Ages
Years
Actual Age Distribution
453 Normal Children
4 5 6 7 8 9 10 11 12 13 14 15 Years

Figure 1

TABLE 8

Test with Delinquents

(Cunningham & Porteus, 1915)

Case No.	Age	Porteus Test	Deficiency	Personal Records
1	13.3	10 years	—3	Truancy and incorrigible.
2	12	10 "	—1	Orphan, committed as neglected child.
3	15.9	11 "	—2	Committed as a neglected child.
4	12.3	10 "	—2	Truant and incorrigible.
5	14	10 "	—3	Neglected child.
6	16.5	11 "	—2	Illegitimate. Committed to Home.
7	12.7	12 "	— ½	Neglected child.
8	12.7	9 "	—3½	Committed to Home as neglected child.
9	13.10	10 "	—3	Neglected.
10	11.9	12 "	0	Neglected.
11	11	12 "	—1	Father died, mother unable to support.
12	11	10 "	—1	Neglected.
13	13½	12 "	—1	Neglected.
14	13½	9 "	—4	Truant, petty thief.
15	11½	9 "	—2½	Truant, thief.
16	12½	8 "	—4½	Illegitimate. Petty thieving.
17	11	9 "	—2	Neglected.
18	12.3	10 "	—2	Neglected.
19	14.2	11 "	—2	Bad companions. Petty thief.
20	12.3	11 "	—1	Truant.
21	11.10	11 "	—1	Truant.
22	11.7	9 "	—2½	Petty thief.
23	14	8 "	—5	Larceny. Absconded twice from Reformatory.
24	16	11 "	—2	Housebreaking, absconded 3 times.
25	13	10 "	—3	Shop-breaking. General bad character.
26	16.3	12 "	—1	Larceny, stealing from shops.
27	15.6	11 "	—2	Unlawful assault.
28	17.4	11 "	—2	Housebreaking.
29	17	9 "	—4	Vagrant.
30	16	10 "	—3	Vagrant.
31	16	7 "	—6	Vagrant.
32	16	12 "	—1	Vagrant.
33	15	9 "	—4	Serious assault. Obscene language.
34	14.4	12 "	—1	Vagrant.
35	16.10	11 "	—2	Thief. Incorrigible. Stealing from houses.
36	17	9 "	—4	Vagrant.
37	15.6	9½ "	—3½	Vagrant and thief. Absconded 5 times.
38	16	8 "	—5	Larceny. Serious assault. Bad character.
39	14.6	8 "	—5	Burglary. (In company with cases 24 and 28.)
40	15	8 "	—5	Vagrant. Stealing letters. (Specific disease.)
41	15	9 "	—4	Murdered companion.
42	14.8	8 "	—5	Vagrant.

TABLE 8—*Continued*

Tests with Delinquents (Girls)

Case No.	Age	Porteus Test	Deficiency	Dr. G. Halley Personal Records and Remarks
43	16	7½	5½ years	Immoral.
44	14	8	5 "	Unemployable. Vagrant. Mental Defective.
45	19	10	3 "	Violent temper. Uncontrollable at home.
46	18	10	3 "	Immoral.
47	18	7½	5½ "	Illiterate. Immoral.
48	18½	6½	6½ "	Immoral. Unemployable.
49	18	5	8 "	Illiterate. Immoral. Unemployable.
50	15	10	3 "	Neglected. Committed to Home.
51	17½	10	3 "	Immoral. Vagrant.
52	16	5	8 "	Mental defective. Unemployable.
53	19	5	8 "	Mental defective. Neglected.
54	16	10	3 "	Vagrant. Runs away.
55	17	12	1 "	Fair mentally. Neglected.
56	18½	13	0 "	Good worker. Family history bad.

revision. The effect of even small alterations either in the direction of lessening or of increasing the difficulty of the maze designs could thus be estimated.

Application of the Test to Delinquents

In order to show the applicability of the tests to juvenile delinquents a summary of results for 56 cases is given in Table 8. Unfortunately, the Binet results are not available for comparison. A brief entry regarding the personal history of each case follows the test records.

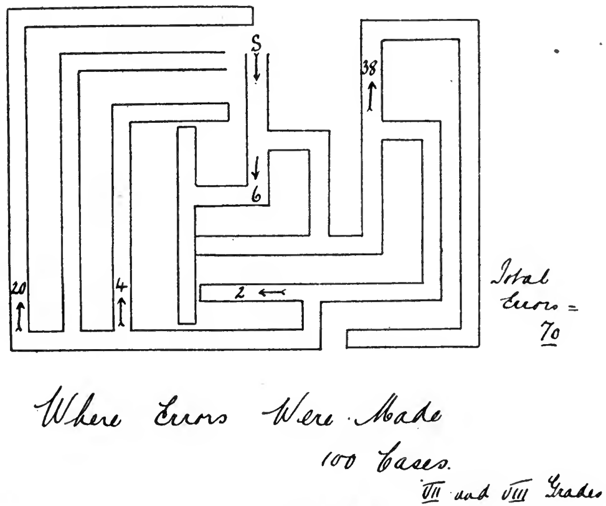
The expectation that the habits of impulsive and ill-considerate action characteristic of this class of children would be reflected in the test-results is amply fulfilled. It was very rare, indeed, to find an individual who succeeded in the 13-year test. The majority tested far below their chronological age, even though no definite diagnosis of mental deficiency had been made previously.

In this and the following investigation the numbers tested are small. The results are reported merely as indicating directions in which the tests might be usefully applied.

Cases 1–22 were from the Burwood Boys' Home, cases 23–42 from the Royal Park Reformatory, cases 43–56 from the Home for Girls, Burra, South Australia.

In Table 8 the fourth column gives the difference between the age by the test and 13 years—the highest test of the series. Considering the fact that the chronological ages are so much

Ten Year Test.

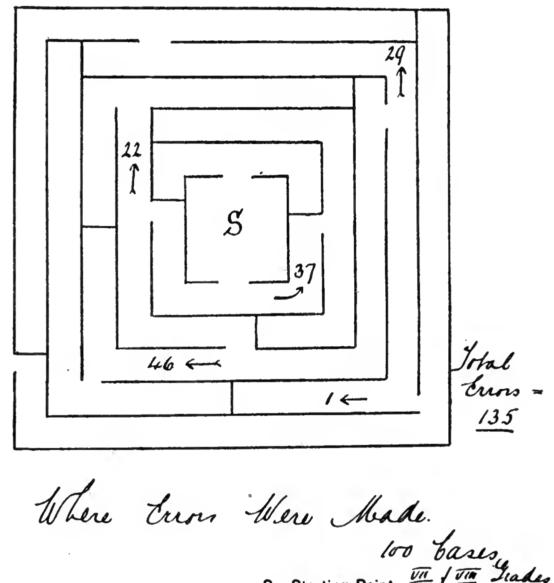


S.—Starting Point.

⟶.—Direction in which to move.

FIGURE 2

Twelve Year Test.

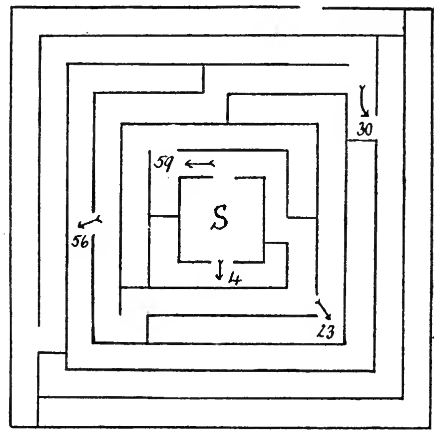


S.—Starting Point.

⟶.—Direction in which to move.

FIGURE 3

Thirteen Year Test.



Total Errors = 172

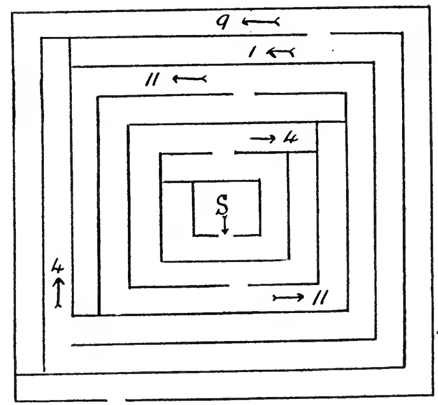
Where Errors Were Made.
100 Cases.
VII / VIII Grades.

S.—Starting Point.

→.—Direction in which to move.

FIGURE 4

Eleven Year Test.



Total Errors = 40

Where Errors Were Made.
100 Cases.
VII / VIII Grades

S.—Starting Point.

→.—Direction in which to move.

FIGURE 5

above the latter limit, the deficiency is really much greater than appears. All the cases listed have been committed by the Courts to the care of the Homes or to the Reformatory. Some were more neglected than delinquent, but in every case the child's character or association rendered commitment to the Home necessary. Group 1 represents criminals in the making, Group II (cases 23–42) the criminal already made. Group III (cases 43–56) is made up of cases who must for their own protection, be kept under restraint. These latter results are available through the courtesy of Dr. Gertrude Halley, Schools' Medical Officer, South Australia.

Deaf and Dumb Children

A small group of children was examined at the School for Deaf and Dumb, Melbourne.

In applying the test, the procedure adopted was to illustrate the task by a test at a much lower level than the child was expected to reach. When the child understood what was required, he began with the next higher test. In the case of the lower tests the procedure was illustrated by a test of similar design.

With the exception of certain children who were proved mentally deficient, both by the experience of the teachers and the verdict of the tests, these subjects did not differ markedly in their performance from other children. The slow and methodical disposition was not often seen. Perceptions were apparently acute, the errors being made chiefly through inpulsiveness rather than through irresolution. The puzzle element in the tests appealed to them greatly. A small group was tested merely to prove the applicability of the tests to these cases.

TABLE 9

Deaf and Dumb Children

Case No.	Age	Porteus Test	Difference	Case No.	Age	Porteus Test	Difference
1	15 years	13		13	12	13	+1
2	12 "	12	0	14	9	9	0
3	14 "	13		15	9*	6	—3
4	15 "	13		16	9*	5	—4
5	13 "	13		17	10	9	—1
6	11 "	10	—1	18	9	8	—1
7	11 "	11	0	19	9	10	+1
8	12 "	13	+1	20	14	12	—1
9	15 "	10	—3	21	15	12	—1
10	12 "	13	+1	22	15*	7	—6
11	13 "	12	—1	23	13*	9	—4
12	12 "	10½	—1½	24	12*	5	—7

* Mentally deficient.

Summary of Conclusions

1. The tests, with an adjustment of the scoring conditions, may be applied with advantage to normal children as well as to mentally deficient. The adjustment is made by taking into account the number of trials required by the child in passing the tests.

2. By careful observation on the examiner's part, facts regarding children's dispositions become apparent through the use of the tests. These facts are not revealed by the Binet examination and have an important bearing on the evaluation of intelligence. It is claimed that the tests form a necessary supplement to the Binet scale.

3. Failures in the tests are mainly due to impulsiveness, habitual carelessness, irresolution. These operate alike in the normal and in the mentally deficient, in the latter to a greater degree. Mental deficiency appears to be merely an extraordinary mental dullness, and it is the writer's opinion that no single test can be devised which will be absolutely diagnostic. Borderline cases must be decided by the application of a number of tests.

4. The relation between this series and the Binet tests is, for large groups of children, fairly constant: about 70 per cent. test within one year of their Binet age.

5. Girls' performances on the whole differ distinctly from that of the boys. The former's results tend to group themselves closer around the Binet ages with more below than above.

In the case of boys the position is reversed. Possible explanations of the differences are:

(*a*) Chance, due to an insufficient number tested, though there is fairly conclusive evidence against this theory.

(*b*) Actual sex differences, *e. g.*, greater tendency to impulsive behavior on the part of girls, over-anxiety, nervousness, lack of planning capacity.

6. The tests may be re-applied with advantage after one or two years interval. Development of the capacities tried out by the tests does not proceed evenly, chronologically speaking.

7. Interesting results may be obtained by the application of the tests to delinquents. Characteristic behavior is reflected in their working of the tests.

8. They may be applied also to deaf and dumb children, and by their aid an estimate of intelligence and educability may be obtained.

FRESHMAN TESTS AT THE STATE UNIVERSITY OF IOWA

IRVING KING AND JAMES M'CRORY

State University of Iowa

In the autumn of 1916 practically all the freshmen entering the College of Liberal Arts at the State University of Iowa were given a series of mental tests. There were 276 women and 268 men.

The tests selected were the following:

1. The Courtis Standard Arithmetic Test, Series B, covering problems in the four fundamentals. This test was given exactly as prescribed by Courtis. A single grade for speed was obtained by adding together all the "attempts." The total number of "rights" was divided by the number of "attempts" for the total percentage of accuracy.

2. A list of twenty-four "mixed relations" or analogies of a fairly difficult type (a set designed by Whipple).

3. Two tests of "opposites." The stimulus words were four hard lists used by Simpson and reported in his "Correlation of Mental Abilities" (1911). These words were rearranged into two lists of forty words each, in order from the easiest to the most difficult, according to the standard of difficulty determined by Gold and King, and published in the JOURNAL OF EDUCATIONAL PSYCHOLOGY, October, 1916.

4. The first three paragraphs of the Completion Test (Mutilated Prose) used by Simpson.

5. Visualization; the three inch cube test, each of the nine questions being weighted on the basis of previous experimentation and assigned a percentage value.

6. Range of Information; Whipple's one hundred technical words. The subjects marked the list to indicate first, which words they could define or explain, and secondly those words with which they were roughly familiar. They were then asked to write the definition or explanation of the first five words which they had indicated as being able to explain. On the basis of these definitions the grade for the entire test was com-

puted, one-fifth of the entire grade being deducted for each wrong definition.

7. Logical Memory, using the paragraph given in Whipple's Manual entitled "The Dutch Homestead." The paragraph was read to the subjects and they were then asked to write as much as they could recall. The papers were graded on the basis of the number of ideas they were able to reproduce according to the standard given in the above mentioned Manual.

All of these tests were given to the Freshmen in groups of from ten to twenty-five. They were timed with a stop watch in each test except in the last three as listed above. The grades for "Opposites," "Analogies," and "Completion" were computed by first marking the errors on each paper and then adding to the time score in seconds, 5 percent. of the time for each error. For instance if it took a student three minutes and twenty seconds to complete a test his score would be 200 which is the number of seconds. If he made on error it would add 5 percent. of that which would make his score 210.

Scoring of Results

In any investigation of this sort, where there are no standards established, it becomes necessary to set up arbitrary standards for grading. This was found to be the case with nearly every one of the tests given. There is a basis, however, for each method used.

Since each of the different tests had a system distinctly its own, there had to be a common basis so that the average of all the tests could be computed and so that the tests could be compared with each other. The following system was devised for such comparison. The median of each distribution of grades was found for each sex; the deviation of each individual from that median was next found; that deviation was multiplied by a factor obtained by dividing the median into 100. This method gave us the grades as multiples of 100. Zero was, of course, the middle case. The algebraic sum of these corrected deviations was then found and divided by the number of tests, which gave the average for all the mental tests. These averages were half above zero and half below, or in other words they were half positive and half negative.

Distribution of Medians and Ranges in the Various Tests

Test	Med. Girls	Med. Boys	% Boys Reaching Med. of Girls	Range Girls	Range Boys
Arith. Speed	50	52	60	19–97	23–97
Arith. Accuracy	74	77	55.4	16–100	24–100
Completion	335	378	36	770–159	841–154
Analogies	228	213	53	480–105	720–85
Opposites Easier	350	384	41	840–144	1128–171
Opposites Harder	900	823	64	1512–318	1538–263
Opposites Comb.	1239	1215	54.1	2055–486	2658–551
Information	25	28	61.6	0–59	0–78
Logical Memory	43	36	32	21–78	9–65
Visual Imagery	33.5	42.5	68	0–100	0–100
Ages	19.62	19.61	50	16–30	17–25

In examining the third column of figures we observe that the boys slightly excel in several of the tests, the exceptions being the completion, easier opposites and logical memory. When the two opposites tests were combined, we find that more than 50 percent. of the boys reach or excel the median grade of the girls. In visual imagery test the boys excel the girls in the greatest degree. We note, however, that the median score made by the girls was only 33.5 out of a possible 100. In this test there were no girls who made a score of zero while there were eight who made a perfect score of 100. Using the median of the girls' scores as a basis and taking the gain of the median of the boys' score over this, we find that the boys accomplished 27 percent. more than the girls.

Variability

Test	Boys	Girls
Arithmetic—		
a. speed	.17	.22
b. accuracy	.17	.16
Completion	.20	.18
Opposites	.13	.18
Analogies	.18	.23
Logical Memory	.22	.17
Information	.25	.32
Visual Imagery	.41	.37

We see from these results that the statement that boys are more variable than girls is not borne out as far as these tests are concerned. If we should add the total variation we find that the boys have 173 and the girls 183 which would make the boys less variable if that one figure were to be taken as a measure. When we divide the tests we find the boys are more variable in four and the girls more variable in the other four.

The above coefficients of variability were computed by dividing the standard deviation by the median.

Discussion of Results

Arithmetic

The Courtis Arithmetic Test, Series B, which we used, has been given to more than a million school children. It is a test in the four fundamentals, and each of the four is usually graded separately. The writers pooled the four grades into a single index. The total number of attempts was taken as speed and the total number of right answers divided by the total number of attempts was taken as the index of accuracy.

Correlations

	Girls	Boys		Girls	Boys
Speed with Accuracy	41	38	Accuracy with Speed	41	38
Completion	22	25	Completion	16	13
Opposites	03	Neg.	Opposites	01	Neg.
Analogies	17	04	Analogies	22	18
Information	06	Neg.	Information	13	Neg.
Vis. Imagery	13	Neg.	Vis. Imagery	20	Neg.
Log. Memory	11	Neg.	Log. Memory	13	Neg.
Test Average	51	56	Test Average	48	41
Uni. Grades	31	36	Uni. Grades	40	36

The above coefficients are all Pearson "r's". They will be used throughout the article.

The writers believe this test should be omitted from any series of tests designed for use with university freshmen. In the first place it takes too much time. Twenty-six minutes were given to this test, which means that by the time the instructions were given, practically thirty-five minutes of the two hours had been taken by this one test. Not only does it take too much time to give, but it also requires a great amount of time to score. There were 97 problems in the four processes and although the answer card is made as convenient as possible, nevertheless it is bound to take more time than the results warrant in this particular kind of testing. Another reason is that given by some of our critics, namely that it tests one particular function under conditions which are not normal. The writers feel that some arithmetic test should be given and possibly the Woody scale in addition would satisfy the conditions. This could be easily graded as it is a scale and not a test. It is what McCall would term a power test and not a speed test. We feel that the reason for including some kind of arithmetic test might be termed a social one. We find in reading the newspapers, magazines, books, or any current literature, that a certain amount

of mathematical knowledge is needed. Not mathematics as such, but the interpretation of statements mathematically given. Addition, of course, does not take care of all these cases and indeed it would be difficult to even list all of them, but since we cannot give a long test to cover all the cases we feel that addition will come the nearest of any of the processes to fulfilling the conditions.

The highest correlation we were able to get between arithmetical speed and any other one test was .25. Bell, at the University of Texas, got an inter-test correlation of .20 between addition and alternatives. Simpson got a raw correlation of .79 between adding and hard opposites. King, at the University of Iowa, tested 56 freshmen engineers. He got a correlation of .75 for attempts or speed and scholastic rank. He got .48 for accuracy and scholastic rank. McCall used 88 6B pupils in New York City. He got an "r" of .20 for adding and school mark. His "r" for arithmetic was .30. His arithmetic test consisted of six selected examples, but details as to their precise nature are lacking.

Completion

We find the completion test one of the two tests in which less than 50 percent. of the boys reached or excelled the median grade of the girls. We also find the boys grades distributed over a wider range than the girls.

Correlations

	Girls	Boys
Completion with Speed	.22	.25
Accuracy	.16	.13
Opposites	.31	.79
Analogies	.46	.58
Information	.21	Neg.
Vis. Imagery	.06	Neg.
Log. Memory	.26	Neg.
Test Average	.60	.77
Uni. Grades	.22	.41

The lowest correlation among the tests for the girls was with visual imagery and we also find a negative correlation for the boys. The P. E. for these coefficients is about .026. As far as these coefficients can be used, it would seem that for a single test the completion is better adapted to the boys. Although they have three negative correlations with the other tests, we find they have a higher coefficient both with the average of all the tests and the average university marks.

McCall in *School and Society* for Jan. 6th, 1917, gives some coefficients which are of interest in this connection. His article is entitled, "Correlation of some psychological and educational measurements, with special attention to the measurement of mental ability." One of his conclusions is as follows: "Meaning by mental ability a composite of all the measurements, the omnibus and completion tests correlate with it 1. and .96 respectively. That is to say, a perfect measure of an individual by omnibus or completion would be a substantially true index of his mental ability." McCall's subjects were 88 6B grade public school pupils in a typical elementary school in New York City. For his completion test he used the Trabue language scale. McCall got a coefficient of .21 with arithmetic (adding), while our coefficient with Series B is practically the same, .22 and .16 for speed and accuracy respectively for the girls and .25 and .13 for the boys. His coefficient was .52 for teachers' marks while ours was .22 for the girls and .41 for the boys. The difference in the age and mental attitude of these two groups must be taken into consideration when any comparison is made. From our results we are not ready to recommend the completion test for use in judging the university student in as thorough-going fashion as McCall does for the sixth grade pupil when he says "The completion test will be a substantially perfect measure of his mental ability."

Bell found in his study of the freshmen at the University of Texas that the completion test showed the highest correlation of any of the test scores with the class marks. The writers found that for the boys two other tests correlated higher than the completion tests with class marks and two others practically the same, while for the girls, the coefficient for completion was next the lowest of any of the tests with class marks.

Dr. King found when he tested a group of 16 Juniors in the College of Applied Science and correlated the results with the combined judgments of three instructors in that department, that he got a coefficient of .35 between the completion test and these combined judgments. The judgments were given simply as to the ability of the men in question.

The writers' results show no important sex differences in the completion test. No reason is assigned for the girls making the better score, on the whole, than the boys.

Opposites

In the easier opposites we find only 41 percent. of the boys reaching the median for the girls, but in the harder opposites we find that 64 percent. reached the median for the girls. In order to obtain a single index for the opposites we added the scores made in each test and called that the score for the opposites combined. When we distribute those scores we find 54.1 percent. of the boys reaching the median for the girls.

Correlation

	Girls	Boys
Opposites with		
Speed	.03	Neg.
Accuracy	.01	Neg.
Completion	.31	.79
Analogies	.52	.77
Information	.24	.56
Visual Imagery	.07	.56
Logical Memory	.32	.38
Test Average	.51	.88
University Grades	.45	.84

We see from the above coefficients that the opposites correlate rather highly with all the other tests for the boys, but not so highly for the girls. We do find, however, that the opposites correlate the most highly of any of the tests *with university marks* for both boys and girls.

We believe that in giving both the easier and harder lists of opposites we obtain a measure which would not be possible by using only one of the tests. Following is the opening paragraph of an article in the JOURNAL OF EDUCATIONAL PSYCHOLOGY for October, 1916, by Irving King and Hugo Gold, of the University of Iowa.

"Among the various mental tests the test of ability to give opposites has found much favor with a number of investigators. Such tests are supposed, if the words are easy, to measure, in some degree, the rapidity of controlled associations reactions; if the words are hard, including such words as "unless," "suave," "ignorant," etc., the test measures in some degree logical keenness in selecting the word which will express most nearly the contrary idea."

Simpson says that the easy opposites is a good test for *readiness for controlled associations.* King got an "r" of .24 between hard opposites (two of Simpson's lists of 20 words each) and

ability of the students as ranked by the three instructors mentioned before in this article. In the case of the 56 freshmen engineers King got a coefficient of .26 for hard opposites and scholastic rank.

Waugh at Beloit College found a correlation of .84 between class standing and opposites.

Kitson, at the University of Chicago, found a correlation of .84 and .86 for the easy and hard opposites respectively, when correlated with the combined index of all his tests. These coefficients were obtained by the product-moments method. The coefficient, using the method of rank differences, as the writers did, was .53 when correlated with standings in the net score.

Kitson's word lists were very easy, his hard list being much easier than the writers' easy list. He defends the use of these easy lists by attempting to eliminate all words that are ambiguous. According to his opinion "these ambiguous words are productive of long pauses and incorrect responses because of conflict of impulses." This, it seems to the writers, is one of the things we are trying to test. Surely one who is able to make a quick, accurate, selective response has better mental ability than one who allows his mental attitude to be upset every time he is called upon to make such a decision. According to Kitson's results, 21 out of 40 made a perfect score of 100 percent. in the hard list, while 31 out of 40 made a perfect score in the easy list. Can we really compare these people when so many of them make a perfect score? If we are trying to determine the percentage of accuracy in that manner, then it seems it would be possible to select a still easier list or in other words, one of zero difficulty.

The writers take practically the opposite stand when selecting words for the opposites test. The answer card which we used was standardized to such an extent that it was fairly indicative of what the majority of the answers were likely to be. We feel that if we were to choose any *one* of our tests as a sign of the mental ability of the individual being tested, it would be this combination of the two opposites tests.

Logical Memory

Memory, perhaps, was the first function of the mind to undergo a test. Nearly everyone who has attempted to make a test

series has had in it some test which had to do primarily with memory. We used for our memory test a passage which was intended to test logical memory or memory for ideas in their logical order.

Correlations

	Girls	Boys
Logical Memory with		
Speed	.11	Neg.
Accuracy	.13	Neg.
Completion	.26	Neg.
Opposites	.32	.38
Analogies	.32	Neg.
Information	.31	.13
Visual Imagery	.31	.12
Test Average	.53	.65
University Grades	.35	.40

Simpson found a correlation between estimated intelligence and memory of words of .93 and memory of passages of .35.

Kitson found a correlation of .29 between logical memory and the net score of all the tests. This was for the forty students in the College of Commerce in the University of Chicago.

In the group of Junior engineers mentioned elsewhere, King found a correlation of .17 between logical memory and estimated intelligence. He found a correlation of .20 between logical memory and completion while ours was .26 for the girls and slightly negative for the boys. He found a correlation of .27 for logical memory and opposites, while ours was .32 for the girls and .38 for the boys.

The writers feel that some test of memory should be incorporated in any series of tests. There does not seem to be sufficient difference between the results of immediate and deferred memory to warrant taking the time to give both tests. If these tests are to be practical at all they must have the so-called mechanical features refined to such a degree that it will minimize the effort and economize the time of those concerned in both the giving and the taking of the tests. The test for immediate memory would be better than the deferred as far as the time factor is concerned.

Visual Imagery

We find this test to be the one in which the boys excel the girls score to the greatest degree. We find the range for each of the groups to be the same, namely, 0 to 100. There were

nine in each group that made zero. There were eight girls who made a perfect score while only five boys made a perfect score. We do find, however, that 68 percent. of the boys reached or excelled the median made by the girls.

Correlations

	Girls	Boys
Visual Imagery with		
Speed	.13	Neg.
Accuracy	.20	Neg.
Opposites	.07	.56
Completion	.06	Neg.
Analogies	.20	Neg.
Information	.23	Neg.
Logical Memory	.31	.11
Test Average	.72	.62
University Grades	.32	.21

Analogies

W. S. Miller, director of the university high school at the University of Minnesota, has probably done more than any other one person in using the analogies test. He gets a correlation of .79 with the analogies and class standing. He gave the tests individually, however, while ours were given in groups.

Correlations

	Girls	Boys
Analogies with		
Speed	.17	.04
Accuracy	.22	.18
Completion	.46	.58
Opposites	.52	.77
Visual Imagery	.20	Neg.
Information	.28	Neg.
Logical Memory	.32	Neg.
Test Average	.72	.74
University Grades	.14	.40

This is a test in logical relation. For some reason it has not found favor with those who have attempted to test college freshmen. The Binet-Simon tests for general intelligence and the Yerkes Bridges point scale use it, though only for adult measurements. It seems to the writers that these tests measure something that is not found in other tests used in this series, and that it would be a mistake not to include them in a series of tests intended for use with college freshmen.

Range of Information

We find 61.6 percent. of the boys reaching the median of the girls in this test. We also find a wider range among the boys,

the highest grade being 78, while the highest grade of the girls was 59. This test, as we gave it and scored it, is at best a sort of a guess affair, the only check being the definitions which the student was asked to write on the back of the sheet. It would be perfectly possible for the student to check each of the 100 terms and then happen to know the meaning of the first five, in which case he would get 100 percent. This case, while possible, is not very probable as is seen when we look at the median performance of these people and find it was only 25 words out of 100 for the girls and 28 for the boys. Unfortunately no one else has published any results of this test computed as ours were so that figures for comparison are not available.

Correlations

	Girls	Boys
Information with		
Speed	.06	Neg.
Accuracy	.13	Neg.
Completion	.21	Neg.
Opposites	.24	.56
Analogies	.28	Neg.
Visual Imagery	.23	Neg.
Logical Memory	.31	.13
Test Average	.60	.65
University Grades	.41	.44

The chief value of these tests lies in the supplementary light they cast in disciplinary cases. By means of these results the Dean of Women and the Adviser of Men were able to determine whether a student failing in his work, was failing from lack of mental ability or for other causes. Below will be given the records made by certain individuals. These are not typical cases, but are selected from among the best and the poorest records in the *mental tests*. It is noticeable that they have good and poor records in their academic work as well.

Case No. 1.

This is a boy who made the best record in his University work. He was the only member of his class of some 500 students who made four "A's." He is carrying English, German, Mathematics, and Zoology. He was below the median in the mental tests in only one test that being speed in arithmetic. He had an accuracy of 88 percent., but he attempted only 51 out of a possible 97. *His rank in his academic work was first while in mental tests it was sixth.*

Case No. 2. This is the record of a boy who had next to the best record in both mental tests and in his academic work. He received two "A's" and two "B's," his course being English, German, Mathematics, and Economics. He was above the median in all the mental tests and was able to complete all the 97 examples in the arithmetic test, being one of three in both the boys' and girls' groups to complete the entire list.

Case No. 3. This is the record of a boy who ranked first in mental tests and who received next to the highest grade in his academic work. He ranked first in the tests, although he attempted only 45 of the 97 examples in arithmetic and was only 65 percent. accurate.

Case No. 4. This is the record of a boy who made zero for his university work, although he took only two of the four examinations. He ranked 259 in mental tests out of 268 boys. He was above the median in only one test, that being range of information. He is a boy who was classed by those who examined him most carefully as below normal. He was in the hospital for some time, having adenoids and tonsils removed. Later a hernia operation was performed. This boy left home when he was 15. His father is a veterinary surgeon and from all indications the home environment was good. When he left home he went west, then served in the navy for four years. During that time he went around the world and naturally acquired a great deal of information the average boy of his age would not have. He was advised not to re-register for the second semester.

Case No. 5 is a boy who ranked 267th out of a possible 268 in mental tests and 234th in academic work. He also was asked not to re-register. The trouble with his academic work was ostensibly ill health, as he had 26 absences excused on that account. He came to the university with very poor preparation, and showed little resistance to distracting influences. He was a constant cigarette smoker, and was consistently reported by his professors as on the delinquent list.

Case No. 6 was a boy who ranked 258th in mental tests and 214th in his academic work. He was utterly lacking in initiative, had no ambition, and failed in all his academic work.

Case No. 7 was a boy who ranked 245th in the tests and 233d in academic work. His trouble, perhaps, is poor preparation,

as he is the fourth student to come to the university from his high school who had to be sent home on account of poor work. He was reported as delinquent from the very first. He was overwhelmed and showed no ability to adjust himself to the situation. He was asked not to re-register for the second semester.

Case No. 8 was a boy who ranked 253d in the tests and 234th in academic work. There was a distinct lack of preparation.

We were not able to obtain as much personal information about the girls as about the boys. The tests tend to pick them out, however, quite as accurately as the boys.

Case No. 9 is a girl who took first rank in the mental tests and also first in her academic work. There were seven girls who made the same grades in academic work, *viz.*, two "A's" and two "B's." This girl is seventeen years of age.

Case No. 10 is a girl who ranked second in mental tests and made the best grade in her academic work. She is sixteen, the youngest member of the freshman class. She has never been reported on the delinquent list and her work is of a generally high order.

Case No. 11 is a girl who ranked fourth in mental tests and received next to the highest grade in her academic work. She is an exceptionally strong student and has never been reported on the delinquent list.

Case No. 12 is a married woman who was the oldest member of her class. She ranked 276th in mental tests, the lowest rank. She left on account of illness before the end of the semester, but was doing failing work when she left.

Case No. 13 is a girl 20 years old who ranked 268th in mental tests. She failed in all her academic work and was asked not to register for the second semester. She was utterly lacking in ability, as she had tried the work the previous year and failed.

Case No. 14 is a girl who ranked 273d in mental tests, and failed in all her academic work. She also was doing the work the second time, and had been reported as failing from the start. She had tried a private tutor to help her with her work, but was simply lacking in ability.

Students Who Were Asked Not to Re-register

There were thirteen girls and fifteen boys who were asked not to re-register for the second semester on account of failing work

during the first semester. In each of these lists we find the median falling in the lower 25 percent. of the ranks in mental tests. The highest rank of those who were asked to leave was a girl who had an average of 13.3 points above the median for her class, and whose rank was 75. This girl failed in more than half of her academic work and was constantly reported to the Dean of Women as on the delinquent list. She was appealed to through her mental test record but to no avail. It was simply a case of too many dates with young men friends, and too many movies. She was perfectly capable, according to the testimony of the Dean, but her attention was on other things.

Distribution of Averages and Ranks of the Boys

	Mental Test Ave.	Mental Test Rank
Case No. I	3.3	111
II	1.2	126
III	.6	129
IV	— 7.6	164
V	—12.1	186
VI	—13.2	191
VII	—13.3	194
VIII	—13.7	199
IX	—15.2	206
X	—17.1	212
XI	—22.7	231
XII	—31.4	245
XIII	—32.0	250
XIV	—35.1	253
XV	—41.0	258

Total Cases 268
Range of Ranks in this Group 111—258
Median Rank 199

Girls

	Mental Test Ave.	Mental Test Rank
Case No. I	13.3	75
II	10.8	83
III	1.3	128
IV	— .1	142
V	— 3.3	162
VI	— 7.9	182
VII	—10.5	193
VIII	—11.2	199
IX	—13.8	208
X	—25.0	242
XI	—37.3	258
XII	—41.8	263
XIII	—45.6	268

Total No. Cases 276
Range of Ranks in this Group 75–268
Median Rank 193

Conclusions

I. This series of tests show fairly good correlations with academic work for the first semester.

II. Inter-test correlations show that we are testing a variety of mental functions.

III. Results show that this series, with the possible exception of the Courtis Arithmetic Test, Series B, is suitable for testing freshmen.

IV. These tests pick out those of poor mental ability much more accurately than those of good mental ability.

V. Certain mental tests would be valuable for colleges to use in dealing with disciplinary cases.

VI. The median performance of the boys slightly excels that of the girls in six of the eight tests.

VII. Sex differences in mental functionings are negligible as far as these mental tests are concerned.

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
LANCASTER, PA. BALTIMORE, MD.

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

The Journal is issuing a series of Monograph Supplements for the publication of more extensive researches in educational psychology. For details prospective contributors may apply to Warwick & York, Inc., Baltimore, Inc.

The price of the Journal is $3.00 per year of ten monthly issues if paid in advance. Back volumes are $3.50 in original paper binding or $4.75 bound in library buckram. Single issues are 50 cents each. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc., 8 West King St., Lancaster, Pa., or 10 E. Centre St., Baltimore, Md.

VOLUME IX JANUARY, 1918 NUMBER 1

EDITORIAL

The War and Educational Measurements

The war is teaching us many things. It is teaching us, who have been the most spendthrift and extravagant people in the world, to economize. It is teaching us that big achievements do not spring into existence from filmy clouds, but have to be carefully prepared and wrought out. It is teaching us the supreme value of knowing how to do things. We have long had the reputation of being a practical people, and we have doubtless gloried unduly in our assumed practical genius. Now we are being brought to appreciate by bitter experience the need for scientific, organized knowledge, for highly trained, skilled direction, for the disciplined imagination of the specialist. But these things are not to be had for the wishing. They are the result of foresight, of careful planning, of a wise education.

We as a people have always proclaimed our belief in education. Our political leaders from Washington onward have seen in education the safety of the nation. But we have never taken the trouble to

scrutinize very closely what that education should mean. Now, under the pressure of war needs we are brought to test, to evaluate, to measure results. There is little doubt that our war experiences will bring about noteworthy changes in education. There will be a sharpening of definition in the aims of education, and a more insistent demand for tangible evidences of results from the methods employed in the process of education. Here lies the opportunity for the advocates of educational measurements. The aims of education are social, but they can and should be expressed in quantitative as well as qualitative terms. The methods of education are incidental to the personal relationship between teacher and pupil, but school procedure should eventuate in measurable attainments. The past decade has seen the development of a considerable number of scales and standards of educational measurement, but these have as yet found small place in the routine of the school room. Teachers and supervisors have been suspicious of the quantitative movement, and have held aloof from active participation in pupil measurements. The time is now opportune for a campaign of public enlightenment in regard to the instruments already at hand for the measurement of school attainments, and to the advantages that will accrue to the superintendent, the principal, the teacher, the pupil, and the public at large from their systematic use.

To this end increased effort will be made in forthcoming issues of the Journal of Educational Psychology to emphasize the practical values of educational scales and measurements. Articles recounting the use of these scales in school room procedure, with resultant diagnosis and treatment of individual difficulties will be especially welcomed. New scales and educational tests will be presented, and the correlation of these with other tests will be encouraged. The measurement of the intelligence of pupils by the Binet and other scales will receive attention as space permits, and studies showing the relation between intelligence ratings and the results of educational tests are invited. Experimental investigations of the learning process will be accepted, and special consideration will be given to those contributions which bear directly upon the work of the class room. In general it will be the policy of the Journal to bring to the attention of superintendents, teachers, and the public the results of such scientific investigations as will increase the efficiency of our schools and make for higher standards of educational achievement.

J. C. B.

NOTES AND NEWS

The Bureau of Educational Measurements and Standards, of the Kansas State Normal School, Dr. Walter S. Monroe, director, has recently issued a series of diagnostic tests in arithmetic which will furnish perhaps the most complete information of a pupil's abilities in arithmetic of any of the tests devised for that purpose. The series contains twenty-one tests, arranged in four parts. Part I, tests 1–6, contains tests of the simpler aspects of the four fundamental operations. There are two addition tests, one single column, three digits, the other four column, five digits each; a subtraction test, a two digit number minus one digit; a multiplication test, a four digit number multiplied by a single digit; and two division tests, one short division, a four digit number divided by a single digit, and one long division, a four digit number divided by a two digit number. They are all speed tests, the time ranging from 30 sec. to 3 min. Part II, tests 7–11, contains a single column addition test of thirteen digits, a four digit by two digit multiplication test, a three and four digit subtraction test, a three by two digit multiplication test arranged with especial reference to the use of zero in multiplication, and a five digit divided by two digit division test. Part III, tests 12–16, presents tests in the addition, subtraction, multiplication and division of common fractions, each answer to be reduced to lowest terms. Part IV, tests 17–21, deals with the multiplication and division of decimals, the digits of the answer being given in each case, and the task being to put the decimal point in the proper place. The total working time for the entire series of twenty-one tests is 35½ minutes. The tests will be furnished to teachers and superintendents at cost, *i. e.*, fifty cents per hundred. If we had a series of tests in denominate numbers and another in percentage, the field of arithmetic would now be fairly well explored.

President-emeritus Charles W. Eliot, of Harvard, speaking on "The America of Tomorrow," asserts that America must be prepared for war for many years to come. He agrees with President Wilson in wanting the establishment of universal democracy, but to secure and maintain this the entire American people must be organized as a great army and navy based on universal training. This would mean a much more highly centralized control of public health, and of efficiency methods in the conduct of education. America must be made to realize how vital it is for the public welfare to have the intellectual productive capacity of each individual developed to its maximum. This involves an intensification and concentration of the learning process in all its aspects.

The application of mental tests to the soldiers in four camps of the national army has yielded such satisfactory results that the War Department has decided to extend the tests to all enlisted men and to all newly appointed officers. It is estimated that this will require a staff of 876 psychological examiners of all ranks for the 31 divisional training camps. In order to provide these (for there are probably not more than 100 psychologically trained examiners available in the entire country at the present time) a school of military psychology will be established at the Medical Officers Training Camp at Fort Oglethorpe, Georgia, in which selected men who have already had some work in psychology will be given intensive training in conducting group and individual psychological examinations, detecting mental incompetents, etc.

The department of psychology of Cornell University calls attention to the courses announced by the University in general, experimental and educational psychology, and in mental tests as offering special preparation for the position of psychological examiner in the army.

The University of Minnesota offers a special war course in methods of psychological examination in accordance with the suggestions of the surgeon general's department. The course is open to men over twenty-one who have had at least one year's work in psychology.

Owing to the fact that both Professor Watson and Professor Dunlap are engaged in war work, instruction in the department of psychology at the Johns Hopkins University has been discontinued for the current year.

Professor Lightner Witmer, of the University of Pennsylvania, has sailed for Europe to take charge of social service work under the direction of the American Red Cross.

Professor John W. Baird, of Clark University, was elected president of the American Psychological Association for the current year, and Professor R. M. Ogden, of Cornell University, and Professor W. F. Dearborn, of Harvard University, were elected members of the Council.

At the Pittsburgh meeting of Section L, Education, of the American Association for the Advancement of Science the following officers were elected: Vice-president, S. A. Courtis, Detroit; Secretary, Bird T. Baldwin, University of Iowa; Member of Council, Guy M. Whipple, University of Illinois; Member of Sectional Committee, Edward F. Buchner, Johns Hopkins University.

The Department of Superintendence of the National Education Association will meet at Atlantic City, February 25 to March 1.

PUBLICATIONS RECEIVED

WILLIAM HOWARD BATSON. *Acquisition of Skill.* Psychological Monographs. Vol. xxi, No. 3. Whole 91, 1916. Pp. 91. $1.00.

This extended study of the learning process grew out of Swift's work with ball-tossing. The experiments included ball-tossing, ball-rolling, picking shot from a rotating table by means of tongs, and releasing shot in a trough at such a moment that they would fall upon a designated portion of a rotating receptacle. Among the conclusions drawn from the experiment are the following: (1) The plateaus of learning depend upon the factors involved in the process to be learned and upon the distribution of attention. (2) The daily fluctuations are common to all types of learning and depend upon the objective and subjective factors involved. (3) The warming up process is the necessary accompaniment of all work where the ordinary daily activity does not provide the special preparation required for performing the work effectively. (4) A short rest period has no marked effect on work one way or the other. (5) After a long rest period the subject is found to be in a condition to improve very rapidly. In some cases the results show that the subjects have actually gained power during the rest period.

COSMOS. *The Basis of Durable Peace. A Series of Papers Written at the Invitation of the New York Times.* New York: Charles Scribner's Sons, 1917. Pp. ix, 144.

A well considered discussion of the conflicts of national interests underlying the war and a plan for an international guarantee of peace and the maintenance of order looking to the prevention of such a catastrophe in the future.

SIDNEY G. FIRMAN. *Progressive Lessons in English.* New York: D. Appleton and Company, 1916. Book I, pp. xiii, 289. Book II, pp. xv, 352.

These books are designed as guides for the language work in the elementary schools, Grades III to VIII. They contain carefully graded exercises for drill in the use of capitals, marks of punctuation, simple parts of speech, business forms, and those verb and pronoun constructions which are the sources of most frequent errors. Perhaps the second book leans too strongly to the illustration of principles in formal grammar to meet the tastes of the radicals in English teaching, but certainly no pupil who has worked through these two volumes, under the guidance of an intelligent teacher, would have any difficulty in avoiding the shocking errors which many of our high school pupils show.

A. P. M. FLEMING AND J. G. PEARCE. *The Principles of Apprentice Training, with Special Reference to the Engineering Industry.* New York: Longmans, Green and Company, 1916. Pp. xiii, 202. $1.20.

Part I deals with the place of the manual worker in industry and the economic importance of training to the individual, to the community, and to the state. In the second part the characteristics demanded from a worker by virtue of a manual occupation are set forth, and the extent to which these are developed by the existing educational system is discussed. The third part emphasizes the need for a means of careful selection of youths according to their vocational fitness and points to the extremely haphazard conditions under which boys at present enter industrial work. The fourth part is devoted to the consideration of artisan training. Existing methods in various industrial countries are described, and the underlying requirements of training stated, followed by the development of a plan suited to modern conditions.

H. W. Foght. *Rural and Agricultural Education at the Panama-Pacific International Exposition.* Bulletin 1916, No. 2. Washington: Bureau of Education, 1917. Pp. 112. 25c.

The bulletin gives a description and photographic illustrations of the exhibits supplied by the Bureau of Education and by the different states. There is also an account of rural education in Alaska and the Philippines and of various agencies that are operating to improve rural education and rural life.

Forsyth Dental Infirmary for Children. Second Annual Report. Boston: 1917. Pp. 31.

Significant in the work of this admirable institution is the establishment last fall of a Training School for Dental Hygienists. The course of study covers a period of twelve months and qualifies the pupils to operate as dental hygienists in the prophylactic treatment of the teeth, and to act as dental nurses in private offices and institutions. The infirmary treated over twenty-nine thousand cases in 1916.

The Fraser Budget for Personal or Family Expenses. New York: Tapley Specialty Company, 1917. Pp. 28. Seventy-five cents.

All cities and corporations are coming to realize the value of a budget, and in these days of enforced economy the same considerations apply to the finances of families and individuals. The author of this scheme has arranged the blanks so that each day's expenses are itemized on a single line, while one can see at a glance the total amount of any one type of expense for the month. At the bottom of the page are spaces for the totals, for the amounts alloted for the various items at the beginning of the month, and for the expenditures of the previous month. Thus one knows exactly where one stands at any time, and can determine where there is a possibility of retrenchment.

FRANK N. FREEMAN. *How Children Learn.* Boston: Houghton Mifflin Company, 1917. Pp. xiv, 322.

In spite of the recent experimental work in educational psychology the teaching of psychology to prospective teachers still remains in a most unsatisfactory condition. Examination questions set by examining boards, and the courses in psychology in normal and training schools still show the predominance of the old common sense, moralizing, faculty psychology, or else the dependence upon such texts of pure psychology as James, Angell, Titchener, or Pillsbury. One reason for this backwardness in the application of educational psychology has been the lack of suitable texts. The present volume will do much to fill this gap. The author's point of view is frankly behavioristic, and the learning process is regarded as the organization of responses. Beginning with a brief account of the nervous system as the basis of response, the author considers the relation of heredity and acquisition in responses, reviewing the work of Galton, Davenport, Woods, Thorndike and others, surveys the chief native responses of the child, analyzes the acquisition of skill, shows how perceptions and associations are built up, develops the principles of memorizing, and illustrates the processes of problem solving or thinking. There are good chapters on the development of language, and on the transfer of training. The book is a distinct advance over anything that is at present available, and deserves extended introduction into normal schools and teachers' reading circles.

JOHN LEONARD HANCOCK. *Studies in Stichomythia.* Chicago: The University of Chicago Press, 1917. Pp. v, 97. Seventy-five cents.

An investigation of the short single verse or double verse dialogue as used in the Greek drama, in Seneca, Mediaeval Latin, Early Italian, French and English, Shakespeare, and in the prose fiction of George Meredith.

WILLIAM HEALY. *Mental Conflicts and Misconduct.* Boston: Little Brown and Company, 1917. Pp. xiii, 330. $2.50.

A study of the causes of delinquency as revealed by an analysis of cases in the Chicago Juvenile Psychopathic Institute. Freud and his school have shown that the origins of mental disturbance frequently lie far back in the history of the individual and have been so repressed as to be quite beyond the patient's cognizance under ordinary conditions. While there is nothing Freudian about this book, there is a consistent effort to discover the causes of anti-social behavior in the previous history of the culprit. This results in uncovering serious and significant mental conflicts which at times cause such mental and moral disturbance that violent and criminal acts are committed

in the effort to secure relief. In these cases just as in the cases of disturbed mental equilibrium the task is first to uncover the cause of the conflict and secondly to reeducate the delinquent so that the mental conflict will not occur. Of course the ideal would be to secure such an education for each boy and girl that these conflicts would never arise. But parents and teachers will need to be much more highly trained in psycho-analysis before such a consummation is even approximated.

Hindu Mind Training, by an Anglo-Saxon mother. New York: Longmans, Green and Company, 1917. Pp. xxiv, 536. $3.50.

This is a curious and interesting book in which an intelligent and widely read mother interprets the Hindu system of mind training on the basis of modern educational psychology. The author has studied western systems of mind training from Rousseau to Montessori and has found none of them satisfactory. More recently she became familiar with the Hindu method, and finds therein much that anticipates the recommendations of leaders in experimental education. With Socrates the author holds that the great problem of education is self knowledge, and by the Hindu system this is made attractive by following the line of greatest connection in preference to the line of least resistance taken by European methods. The author ranges easily from Locke and Descartes to Münsterberg and Frank N. Freeman. Especially interesting is the Hindu version of vocational selection and guidance compared with Münsterberg's *Business Psychology*, and the Hindu moral training considered from the point of view of Holt's *The Freudian Wish*. While the term "mind training" itself suggests a point of view that is now considered antiquated by many psychologists and investigators, the examination of the stories and exercises used in Hindu education from the point of view of Titchener, Hollingworth, Ebbinghaus, and other modern psychologists leads to interesting if not all together convincing results.

F. J. Foakes Jackson. *Social Life in England, 1750 to 1850.* New York: The Macmillan Company, 1916. Pp. ix, 338. $1.50.

This is a series of Lowell Lectures delivered in Boston in March of last year. It is an authoritative account compiled from such sources as John Wesley's diary, the life of the poet Crabbe, the story of Margaret Catchpole, reminiscences of Cambridge, the Creevey Papers, and accounts from the newspapers and periodicals of the time. The chronicle is of especial value in helping one to appreciate the social background of Dickens, Thackeray and Trollope.

Carl Kelsey. *The Physical Basis of Society.* New York: D. Appleton and Company, 1916. Pp. xvi, 406. $2.00.

This is an introduction to sociology by way of the relation of the organism to its environment, the facts of heredity, and the idea of evolution applied to human beings. The book is packed with sig-

nificant data for the understanding of the development of man under the influence of his physical environment. Among the topics treated are: Man in Relation to the Earth; The Struggle for Existence in Animal Life; Man's Control of Natural Forces; Man as a Product of Evolution; The Fundamental Doctrines of Heredity; The Development of Society Along Heredity Lines; Racial Mixtures; Sex Relationships; Social Relations and Population; Institutional Restrictions and Personal Initiative; and the Nature of Human Progress. It is a book that may well be taken as the foundation of one's educational theory and outlook upon life.

J. L. Manahan. *A Bibliography of Educational Surveys and Tests.* University of Virginia Record, Volume II, No. 3, 1916. Pp. xx, 47–92.

This extensively annotated bibliography is in five sections. Section I discusses articles dealing with the fundamental principles of educational measurement; Section II, school surveys; Section III, standard tests in school subjects, including arithmetic, handwriting, spelling, reading, composition and grammar, geography, drawing, algebra, and foreign languages; Section IV, studies in the application of standard tests; and Section V, tests of general intelligence. Teachers who are not familiar with recent educational investigations will find this summary very convenient.

Clarence Whittlesey Mendell. *Latin Sentence Construction.* New Haven, Conn.: Yale University Press, 1917. Pp. x, 214. $1.50.

This is an important contribution not only to Latin syntax, but to the principles of sentence structure in general. In the introductory chapter the general psychological principles are outlined that underlie all sentence connection. These lead directly to the examination of the fundamental relations which language seeks to express. The remainder of the book is devoted to the results of investigations on the means employed by language to express these relations. The analysis discloses the extremely interesting structure of even the simplest expression of ideas.

Edwin F. Northrup. *Laws of Physical Science. A Reference Book.* Philadelphia: J. B. Lippincott Company, 1917. Pp. vii, 210. $2.00.

This is not a text-book in physics, nor is it a philosophical discussion of the constitution of the universe, but it is, as the title implies, a ready reference compendium of the most important laws that have been formulated by physicists. In each case the law is stated as succinctly as possible, brief explanations or formulae are given where these seem desirable, and there is a reference or two to standard works or to monographic literature on the topic. It should be a very helpful handbook to college students of physics.

John Clyde Oswald. *Benjamin Franklin, Printer.* New York: Doubleday, Page and Company, 1917. Pp. xv, 244. $2.00.

This study of the "many sided Franklin" confines itself to Franklin's activities in the printing trade. The book contains many reproductions of title pages, magazine covers, catalogues, and newspapers. One of the most interesting of these is the title page to the first number of *Poor Richard's Almanack.* There is an account of Franklin's early apprenticeship, his efforts to get a start in Philadelphia, his partnerships, his experience as a type founder, and his public life as it was related to printing. There is a reprint of the first American cartoon representing the famous snake divided into thirteen sections with the legend "Join or Die."

Rudolf Pintner and Donald G. Paterson. *A Comparison of Deaf and Hearing Children in Visual Memory for Digits.* Reprinted from the Journal of Experimental Psychology, Volume II, No. 1, February, 1917. Pp. 76–88.

The authors find that deaf children as a group have an abnormally poor memory span due to the lack of auditory experience. Lip readers are superior to manual readers on the average. No significant sex differences were found.

Report of the Commissioner of Education for the Year Ending June 30, 1916. Washington: Bureau of Education, 1916. Volume I, pp. xxvii, 692. Volume II, pp. viii, 663.

The statistics compiled in these volumes show that the number of persons enrolled in educational institutions in the United States is increasing at an average rate of half a million annually. This means that approximately 24 per cent. of the inhabitants of the United States are attending school as compared with 19 per cent. in Great Britain, 17 per cent. in France, 20 per cent. in Germany, and 4 per cent. in Russia. The nation's current educational expenditure is in round numbers one billion per year. The chapter on educational surveys by Prof. Edward F. Buchner, of Johns Hopkins University, gives a good account of recent undertakings in this field and will be of distinct help to many thoughtful school people. Another interesting chapter is that on educational hygiene by Dr. Willard S. Small.

Report of the 1915 Legislature Committee of the State of California on Mental Deficiency and the Proposed Institution for the Care of Feeble-minded and Epileptic Persons. Whittier, California: Whittier State School, 1917. Pp. 90.

This report contains extracts from the writings of Charles B. Davenport and Henry H. Goddard, an abstract of the work of Dr. Samuel C. Kohs on the *Distribution of the Feeble-Minded,* a summary

of the report of the New York Commission to Investigate Provisions for the Mentally Deficient, and articles by Lewis M. Terman and J. Harold Williams on feeble-mindedness in children of school age. The text of the bill to establish the new institution is given and there is a brief history of the movement to combat the evils of mental deficiency through legislative and moral prophylaxis.

Report of the School Survey of School District Number One in the City and County of Denver, Colorado. The School Survey Committee, 1916. Part I. *General Organization and Management,* by Franklin Bobbitt. Pp. 166. Twenty-five cents. Part II. *The Work of the Schools,* by Franklin Bobbitt and Charles H. Judd. Pp. 180. Twenty-five cents. Part III. *The Industrial Survey,* by C. A. Prosser and W. H. Henderson. Pp. 91. Twenty-five cents. Part IV. *The Business Management,* by J. T. Byrne. Pp. 108. Twenty-five cents. Part V. *The Building Situation and Medical Inspection,* by Lewis M. Terman. Pp. 75. Fifteen cents. *Supplemental Report on Organization and Administration,* by Ellwood P. Cubberley. Pp. 21.

The most interesting portion of this survey for the scientific student of education is Parts II, III, and V. In Part II compositions from the fourth to the eighth grades were measured by a scale devised especially for the purpose. The form and speed of handwriting were determined for each school and diagrams are presented to show the growth in high schools and the relation between high and elementary schools. In Part III we find a good discussion of retardation and elimination, of manual training in elementary and high schools, and of the place of drawing and design in vocational education. Part V presents statistics to show the inadequacy of the present school grounds and discusses seating, lighting, blackboards, heating and ventilation, and the general subject of school house-keeping.

John J. Schlicher. *Latin Plays: For Student Performances and Reading.* Boston: Ginn and Company, 1916. Pp. vii, 213. Seventy-five cents.

Many educators at the present time believe that high school pupils should not be encouraged to spend their time on the study of Latin. All will agree, however, that if Latin is taught it should be taught in the most effective, interesting and stimulating manner possible. This little book contains seven Latin plays which are intended to be learned and acted by pupils. Two of the plays are closely connected with Caesar, two with Cicero, one with Virgil and one with Ovid. Interspersed in the dialogue are choruses and songs, the music for which is to be found at the end of the text. The progressive high school teacher of Latin will find this selection of plays of great value in keeping up the enthusiasm of pupils.

Paul Shorey. *The Assault on Humanism.* Boston: The Atlantic Monthly Company, 1917. Pp. 80. Sixty cents.

The publishers speak of this book as "an eloquent defense of the classical tradition against the attacks of President Eliot, Dr. Abraham Flexner, and the advocates of the 'new school.'" That it is clever, witty, facile, that it abounds in literary allusions skillfully handled, that it uncovers and mercilessly flays certain inconsistencies of statement must be admitted. But the author's brilliancy only serves to becloud and obscure the issue instead of clarifying it. For the question is not liberalism versus utilitarianism in education, but rather what constitutes a liberal education in this day and age. The classicists have no monopoly of liberal education. Indeed it is becoming increasingly evident that classicism represents an extremely narrow type of technical education. What is needed is a "new humanism," a return to the ideals of the Greeks and of the early Renaissance, a revaluation of educational values in terms of the present needs of humanity. All education is and must be utilitarian in that it meets some sort of human need. But the classical procedure defeats its own claim of being liberal. In the book much is made of the disciplinary value of classical study, and the medieval faculty psychology is stoutly defended. One cannot restrain a smile at the citation of "one of the best modern psychologies for teachers, the little volume of the eminent English psychologist Lloyd-Morgan," as conclusive authority for the acceptance of the faculties in psychology.

Sixty Second Annual Report of the Board of Education of the City of St. Louis, Missouri. for the year ending June 30, 1916. Pp. 639.

This report gives evidence of the growing demand upon school authorities to take the public into their confidence, and not only to tell them in a general way what is going forward in the schools, but to see to it that there are many things worth telling. The special features of this report are a tribute to the late Susan E. Blow, of Kindergarten fame, a series of graphs showing the average per cent. of failures in the several subjects in four high schools, and a detailed report by Dr. J. E. W. Wallin, director of the psycho-educational clinic, on the prevalence of mental and physical defects in school children. A particularly valuable feature of Dr. Wallin's report is the account of a census of speech defectives among the 89,057 pupils of the St. Louis Schools. There are extensive tabulations of the distribution of various types of speech defects and Dr. Wallin gives some excellent suggestions regarding corrective speech work in the schools.

Jessie M. Tatlock. *Greek and Roman Mythology.* New York: The Century Company, 1917. Pp. xxviii, 372. $1.50.

In this elementary book on mythology the author's aim is not to "write down" to children, but to present a simple, straightforward narrative of gods and heroes, couched in such terms as appeal to the

thoughtful, inquiring mind. The book is profusely illustrated with pictures of statues and other classic representations of Homeric characters. An appendix gives a brief list of poems and dramas based on the myths. It is an attractive book to own.

The Teaching of Government. Report to the American Political Science Association by the Committee on Instruction, Charles Grove Haines, Chairman. New York: The Macmillan Company, 1916. Pp. xi, 284. $1.10.

In spite of the fact that the teaching of citizenship is the defense most frequently offered for universal public education, it is generally admitted that most of the teaching of "Civics" is of a dry, formal, text-book nature, with little consideration or regard for the interests of the pupils. Recognizing and deploring this situation, the American Political Science Association in 1911 appointed a committee to make a careful and thoroughgoing investigation of the entire subject. The present volume presents the results of five years of labor on the part of the committee. Part I makes a survey of recent progress in the teaching of government, recounts various organized efforts to improve the instruction, and sets forth purposes, methods, materials, and devices. Part II deals with advanced civics in secondary schools, with some consideration of the training of teachers in this field. Part III presents suggestions for courses of study, with valuable bibliographies. Part IV discusses the teaching of political science in colleges and universities. In general the committee finds that the subject of government is neglected in most college courses and that the facilities for the training of high school teachers in this subject are sadly inadequate. An appendix of eighty pages summarizes the reports on the teaching of civics in elementary and secondary schools made by state committees in twenty different states. The book is a landmark in the teaching of the subject and will undoubtedly have a wide circulation and influence.

Third Annual Conference on Educational Measurements Held at Indiana University, April 14 and 15, 1916. Bulletin of the Extension Division, Indiana University, Volume II, No. 6, February, 1917. Pp 208 Fifty cents.

This volume contains stimulating addresses by Professor Ellwood P. Cubberley on *The Significance of Educational Measurements, Standard Tests in the Work of School Administration*, and *Measurements Applied to School Financing*. Professor Cubberley believes that the introduction of educational measurements will cause a much greater demand to be made on the superintendent and will call for distinct and adequate professional preparation for the work. The monograph also contains four addresses by Dr. B. R. Buckingham, educational statistician for the State of Wisconsin, in which he discusses *Efficiency Indices and Principles of Scale Derivation, with Special*

Application to Arithmetic, History, Geography, and Grammar. There are reports by Professor H. G. Childs on the *Cost of Instruction in High Schools, Enrollment, Acceleration, Retardation and Percent of Failures in High Schools,* and an important investigation of *The Measurement of Achievement in Algebra.* The latter is perhaps the most extended experimental study of the subject that has yet been published.

EDWARD L. THORNDIKE. *English Composition.* One Hundred Fifty Specimens Arranged for Use in Psychological and Educational Experiments. New York: Teachers College, Columbia University, 1916. Pp. 127.

These specimens of English composition will furnish material to determine the ability of prospective teachers to evaluate compositions; to test for individual, class, and sex differences in complex judgments; to determine the amount of transfer of improvement by practice; to detect constant errors in judgment; and to demonstrate and measure the value of objective scales. The general merit of each composition has been determined on the basis of the Hilegas Scale by a consensus of from 23 to 100 judges. These values are given in two tables at the beginning of the monograph and are only to be referred to as a check on the experimental use of the specimens.

W. TROTTER. *Instincts of the Herd in Peace and War.* New York: The Macmillan Company, 1916. Pp. 213. $1.45.

"The general purpose of this book is to suggest that the science of psychology is not the mass of dreary and indefinite generalities of which it sometimes perhaps seems to be made up; to suggest that, especially when studied in relation with other branches of biology, it is capable of becoming a guide in the actual affairs of life and of giving an understanding of the human mind such as may enable us in a practical and useful way to foretell some of the course of human behavior." The psychology here referred to is rather a type of speculative sociology and rests upon the very questionable concept of instinct. The author's theory of gregariousness is developed in the first seventy pages and the remainder of the book is devoted to an analysis of the social forces involved in the present war. While we may not agree with the author's psychology, the discussion of the English and German social attitudes toward the war makes very interesting reading.

Value of the Classics. Princeton: Princeton University Press, 1917. Pp. vii, 396. $1.50.

The classicists are on the defensive. Their backs are to the wall. They feel the necessity of calling out all their reserves to meet the increasing pressure of educational criticism. Accordingly a great classical rally was held at Princeton last June, and expressions of loyalty and support were secured from as many of the faithful as

were believed to have influence with the public. This volume presents the results. The editor, Dean Andrew F. West, of Princeton, introduces the book with the following words: "In education definite evidence is worth more than theorizing." (This sounds promising. "Definite evidence" is what many of us are ardently seeking.) "This book is chiefly an appeal to facts, and two classes of facts appear in its pages. The first includes . . . the statements of nearly three hundred competent observers representing the leading interests of modern life and including many of the highest names in our land. . . . The second part is statistical. The most pertinent and reliable facts in the records of our schools and colleges, so far as procurable, are here presented and examined." Thus it is seen that the facts of the first class are mere opinions. That opinions are facts —very stubborn facts—the history of civilization amply reveals. But that they constitute "definite evidence" would be accepted chiefly by those whose critical judgment has been warped by the "formal discipline" of classical training. The facts of the second class are limited to a few pages of statistics from the reports of the bureau of education and the college entrance examination board. If this is the best the classicists can do, they would seem to be in a precarious position.

William A. White. *Mechanism of Character Formation: An Introduction to Psychoanalysis.* New York: The Macmillan Company, 1916. Pp. vii, 342. $1.75.

The author emphasizes the genetic approach to the subject of character construction and bases his argument largely on the biological work of Marshall and Loeb. In the next chapter however, he dives into the abyss of the unconscious, and from there on we meet the usual Freudian phraseology of the libido, conflict, symbolism, dream mechanisms, the family romance, and the resolution of the conflict by suggestion and re-education. The book is characterized by much greater sanity than many of the Freudian productions. As Professor Watson has recently pointed out, there is the ever present danger in this type of discussion of laying undue stress on a system of verbal concepts, which seem to mean something to their users, but which lead to a crude, mystical, anti-scientific, vitalistic attitude toward fundamental problems. The movement has done some good in turning the attention of medical men to psychology, but there is great danger that the verbiage of its terminology may lead to a stultification of rigid, accurate, scientific thinking.

Willard Huntington Wright. *The Creative Will. Studies in the Philosophy and Syntax of Aesthetics.* New York: John Lane Company, 1916. Pp. 288. $1.50.

This is a sententious disquisition on the significance of art. Form is the basis of all beauty, and aesthetic form must reflect the form

which is most intimately associated with our sensitivities. The artist cannot go to the science of aesthetics to learn how he may produce beauty. Creation in the arts is the result of some impetus received either from objective nature or from a mental process. The second and most extensive part of the book deals with problems of aesthetics while the third and fourth sections treat of art and the artist, and art and the individual.

VOLUME IX FEBRUARY, 1918 NUMBER 2

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

METHODS AND RESULTS OF A CLASS EXPERIMENT IN LEARNING

WALTER F. DEARBORN AND JOHN M. BREWER
Harvard University

I. THE CONDITIONS OF THE EXPERIMENT

In an article in the first volume of this JOURNAL[1], a series of experiments in learning was suggested for class use. The present article gives the results of one of these experiments—namely, on the effects of practice—as conducted with a class of about thirty-five students in educational psychology. Instead of the digit-letter substitution proposed for the practice series in the former article, the practice consisted in learning to write an old civil war code[2]. For convenience of reference, it is reproduced below (Figure 1). As a test of intelligence, the code is usually used but once, *i. e.*, to transcribe a short phrase or sentence. In the present experiment, long passages of connected prose were translated into the figures of the code in an attempt to increase the speed with which the code could be written. The passages were the same as used in the article above

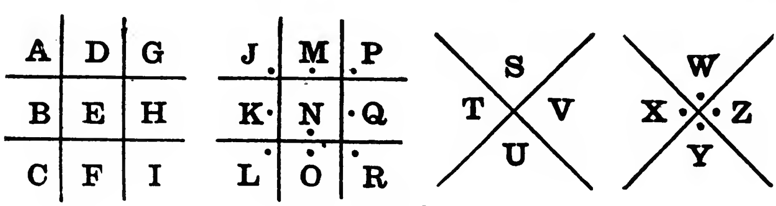


FIGURE 1: CODE.

[1]Vol. 1, pp. 373–384.

[2]This code is described by Healy and Fernald in their "Tests for Practical Mental Classification," Psychological Review Monographs, Volume XIII, No. 2, Test No. XI, and has been used as a test by Goddard, Terman and others, in the various revisions of the Binet-Simon Tests. Compare TERMAN "*Measurement of Intelligence*," Chapter XIX, pp. 330–332.

referred to and the general arrangement of the printed form was similar. The first few lines of a sample practice sheet with the characters filled in is reproduced below (Figure 2). The code was printed on small library cards, and as noted below, was used in one of the preliminary tests. The general principle on which the code was constructed was then explained and the cards collected. In the practice proper the printed code was not used, the subjects being required to reconstruct the code mentally as they wrote,—in exactly the way in which the code was undoubtedly intended to be used originally. The use of the same passages and general arrangement followed in the digit-letter substitution described in the above noted article was for the purpose of comparing the practice curves in these somewhat different sorts of learning.

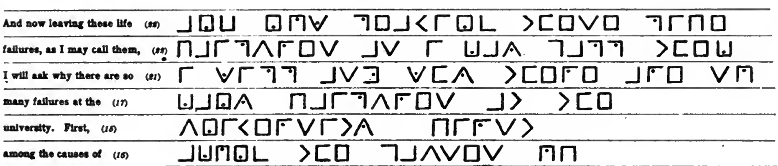


FIGURE 2: PRACTICE MATERIAL

Before and after the practice series, the subjects were given a group of five closely related tests with a view to noting the possible effect of the practice gains in the use of the code on what might appear to be similar activities. Two control groups were given these tests, but without the intervening practice with the code.

The general plan of the experiment may be illustrated by the following diagram.

AN EXPERIMENT IN LEARNING

Practiced Group

1st day	2nd, 3rd....................12th day	13th day
5 tests given	Practice series, with code. 22 trials.	5 tests repeated

Unpracticed or Control Group

1st day		13th day
5 tests given	(no intervening practice)	5 tests repeated

The practiced group was a class in educational psychology. Two control groups were used,—a class in teaching and one in school administration. The five tests to be described below were given to each group of students on two separate occasions, twelve working days apart, and in the interim the class in educational psychology was practiced twenty-two times in the use of a code. (The interval between the giving of the five tests and the repeating of these tests should be the same in all the groups. Circumstances, however, made this impossible in the present experiment. The interval for the practice group was sixteen calendar days; for the class in teaching, nine days; for that in administration, nineteen days.)

The five tests were as follows:

(1) Digit-Symbol Substitution. This test is the same as the one described in the former article and similar to the one described by Whipple in Test 37B of his *Mental and Physical Tests* (revised edition). The printed sheets of paper were distributed. At the top of each sheet are ten circles, containing the ten digits in order, and in each circle is a simple symbol for the digit therein contained. Below are columns of five-place figures. These figures are to be translated into the symbols. The score is the number of substitutions made in two minutes.

(2) Complex Dotting. Sheets of paper were ruled into squares, each square containing either the number 1, 2, or 3, these numbers being distributed at random. The task was to put in each square the designated number of dots, the time allowed being one minute. The score is the number of squares completed. There are one hundred and fifty squares. The first six rows of squares are reproduced in Figure 3.

(3) Number Code. The top of each sheet contained a simple code for the nine digits. Below was a series of sixty characters which the students translated by means of the code. The time taken to complete the translation was recorded[3].

(4) Digit Letter Substitution. Sheets of paper were distributed, each containing at the top circles in which were inserted numbers from 1 to 26. Each circle also contained a letter of the alphabet, the

[3]The test is one used by Brunner and McMillan in the Child Study Departmens of the Chicago Public Schools and is described by Healy and Fernald in the Psychological Review Monographs, Volume XIII, No. 2, Test No. X (Cross Line Test, B).

2	3	1	2	1	1	2	3	2	2
1	1	2	1	3	1	2	3	3	2
3	2	3	2	2	1	3	1	2	3
1	3	2	3	1	1	2	1	3	1
2	1	3	2	1	3	2	3	3	2
1	3	2	1	3	1	2	2	1	1
2	3	1	2	1	1	2	2	1	3

FIG. 3: TEST 2.

twenty-six letters being distributed at random. Under this code were printed sentences, and the task consisted in translating letters into digits for a period of five minutes. These blanks are the same as those used for the practice series in the article above referred to,[4] but with a somewhat simpler text.

(5) Letter Code Translation. The test consists of a sheet of paper containing forty-two simple characters which, by means of a code printed upon separate cards and distributed to the students, can be translated into English sentences. Reproductions of the two code passages used, one before and one after the practice, are given in Figure 4 (a) and (b). The first passage, when translated, reads: "Come quickly. Bring men and supplies for journey. x. y. z." The second passage reads: "Enemy on the march. Meet them at river. Hurry forward. M." The code used is reproduced in Figure 1. The time needed to make the translation was recorded as the score.

The second trials in the above five tests were held in exactly the same way except that new material of equal difficulty was used in the translation and substitution tests.

The practice series was as follows: Using the same code as that employed in test No. 5 above, and reproduced in Figure 1, sheets

[4]This JOURNAL. Vol. I.

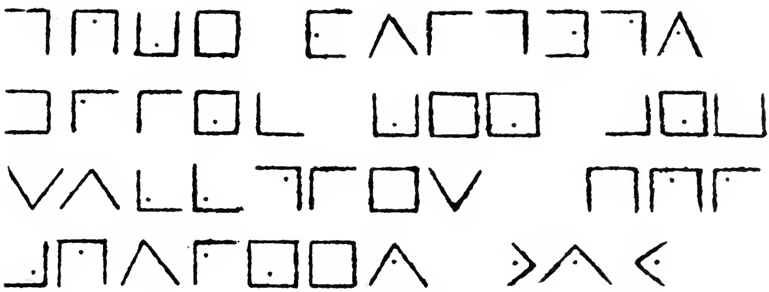


FIGURE 4a: TEST 5.

come quickly
bring men and
supplies for
journey xyz

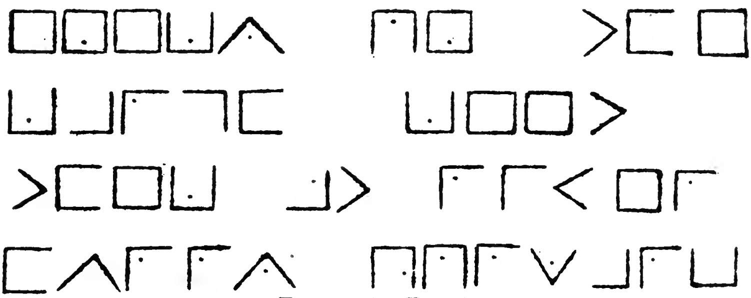


FIGURE 4b: TEST 5.

enemy on the
march meet
them at river
hurry forward
m

of printed material were distributed and the subjects were asked to translate the English sentence into the code. After the scheme on which the code is constructed had been explained, the subjects, as noted above, were required to reconstruct the code mentally instead of using the printed key. Each trial was five minutes long, one trial being made in the morning of each day and one in the afternoon. The first test and the twenty-second test were taken in class; the others at home. Before taking each test the student was asked to record the date, the time of day, his mental and physical condition, and his expectation as to gain or loss. After the test, he recorded the number of letters translated and the gain or loss over the previous performance.

At the conclusion of the experiment each student made up his learning curve on cross section paper. In addition, he made a separate graph of the morning and evening performances and computed his percentage of gain in the practice series—the highest re-

cord compared with the initial performance. Finally, he wrote a discussion upon the following questions:

(1) Discuss the general form of your learning curve.

(2) Compare the morning and night curves.

(3) Were your predictions of any value?

(4) Was the ability gained in practice transferred in such a manner as to make the second record in any of the five trials higher than it would have been without practice?

(5) Was interference noted between the practice series and any of the tests?

(6) Criticize the experiment.

It was the intention, as may be seen from the above description, to have several tests in the preliminary and final series so similar in nature to the exercise employed in the practice that they would presumably be affected either by way of reinforcement or of interference by whatever improvement might be made in the latter. Whether it is a matter chiefly of interference or of reinforcement was left to the subject to decide on the basis of his observations and introspections, although, of course, the scores made needed to be taken account of in the final interpretation of results. In this connection, it is instructive to have the subjects attempt before the experiment to make some prophesy of the probable results.

II. The Results of the Experiment

The following scores were obtained in the preliminary and the final trials of the five tests. These scores are given for the three classes tested. The first tabulation, that for the practiced group, contains also the scores for the initial and the highest performances in the practice series. As the figures represent in three cases the number of substitutions, etc., made in a given time and in the other two tests the time (in seconds) necessary to complete the test, the facts may be restated for convenience of reference.

Test No. 1. Number of substitutions of symbols for digits in two minutes.

Test No. 2. Number of squares dotted in one minute.

Test No. 3. Time (in seconds) necessary to complete number-code test.

Test No. 4. Number of digits substituted for letters in five minutes.

Test No. 5. Time (in seconds) necessary to complete translation of code passage.

TABLE I

Practiced Group

Scores in Tests and in Practice Series

Student	Classroom Tests: Preliminary 1	8	3	4	5	Classroom Tests: Final 1	2	3	4	5	Practice Series: Init.	High	Gain
1	59	106	110	91	160	76	113	85	87	135	49	147	98
2	80	116	90	134	120	82	135	65	162	65	52	244	192
3	51	62	130	87	180	61	87	95	82	90	45	209	164
4	70	128	145	110	155	93	145	75	110	105	30	184	154
5	66	116	95	294	180	72	123	80	294	110	44	279	235
6	97	112	85	168	85	98	122	65	189	65	93	300	207
7	47	78	180	53	180	56	91	115	56	100	22	161	139
8	49	110	105	94	175	67	137	90	111	85	16	237	221
9	53	111	135	85	145	80	121	90	104	105	36	195	159
10	65	110	90	88	115	68	127	80	106	115	95	285	190
11	60	95	75	98	120	68	105	70	130	70	29	205	176
12	46	80	180	77	180	65	96	125	96	170	43	265	222
13	61	115	75	123	100	71	118	85	124	95	66	346	280
14	55	95	95	72	180	67	113	80	106	105	38	293	255
15	20	103	120	123	135	53	116	90	143	145	22	296	274
16	75	90	105	109	110	75	100	70	121	90	52	279	227
17	56	83	75	112	95	79	90	70	122	95	73	252	179
18	66	97	100	101	90	74	111	70	128	75	49	259	210
19	59	100	95	122	180	55	109	85	125	100	26	178	152
20	51	92	120	74	180	49	111	85	76	65	33	221	188
21	55	100	110	98	125	75	113	75	135	110	57	236	179
22	47	101	100	94	180	75	116	80	105	120	30	232	202
23	66	103	180	135	170	87	123	90	192	115	36	278	242
24	53	100	110	109	180	45	102	100	110	145	30	176	146
25	55	110	180	123	145	67	126	130	114	110	50	274	224
26	41	110	80	92	85	47	122	70	95	65	46	217	171
27	56	85	110	121	145	64	88	95	109	155	34	225	191
28	52	98	90	70	135	62	114	80	82	70	24	281	257
29	54	96	115	121	100	67	95	75	142	95	55	282	227
30	65	97	115	99	135	80	109	85	122	60	47	287	240
31	35	111	180	114	180	75	117	80	135	180	40	297	257
32	33	74	180	74	155	47	93	80	92	90	27	190	163
33	70	120	150	80	120	73	122	85	97	100	36	210	174
34	58	103	115	103	165	70	123	75	119	85	20	166	146
35	56	110	75	97	120	66	121	70	103	100	52	171	119
Ave.	57	101	117	107	143	69	113	84	121	102	43	239	196
A. D.	15	11	27	24	22	9	8	11	26	22	14	42	39

TABLE II A

Control Group Number 1

Scores in Classroom Tests

Student	Preliminary 1	2	3	4	5	Final 1	2	3	4	5
1	56	121	100	110	140	71	121	75	135	150
2	59	99	150	173	125	73	105	90	198	85
3	58	68	170	92	155	56	83	110	106	145
4	68	110	110	93	175	70	118	90	123	135
5	81	85	120	118	125	82	107	80	167	105
6	55	108	180	104	180	41	101	165	111	140
7	63	108	120	77	170	65	100	100	139	120
8	88	130	80	101	65	83	136	65	134	90
9	66	102	90	119		73	95	85	138	
10	72	127	105	208	110	69	123	100	220	145
11	75	120	115	54	165	75	132	85	84	165
12	48	100	105	103	165	60	115	95	100	137
13	69	97	130	113	180	63	99	100	143	180
14	75	97	100	164	165	77	112	80	195	130
15	66	93	100	95	150	70	109	85	121	125
16	105	126	65	205	75	104	136	60	219	75

TABLE II B

Control Group Number 2

Scores in Classroom Tests

Student	Preliminary					Final				
	1	2	3	4	5	1	2	3	4	5
1	60	76	105	101	150	74	84	75	128	110
2	50	110	120	113	140	73	120	90	120	95
3										
4	41	87	130	87	180	49	92	110	84	180
5	58	93	90	103	140	64	103	85	96	115
6	27	80	150	77	180	67	103	80	112	130
7	29	70	180	39	95	52	82	155	43	180
8	57	102	105	115	165	66	105	95	135	110
9	71	110	100	106	120	78	110	70	117	85
10	65	72	120	114	85	60	70	95	104	105
11	65	85	130	95	130	75	95	90	129	115
12	33		140	70	115	42		110	63	115
13										
14	55	118	145	67	170	63	127	105	99	155
15	59	102	115	98	105	67	102	80	115	95
16	73	121	90	119	115	85	135	75	130	75
17										
18	65	116	110	97	115	81	120	75	116	95
19	65		125	110	180	80		130	90	140
20										
21	61	100	130	71	170	65	92	95	89	125
Av. (*)	62	101	119	106	141	69	108	93	124	124
A. D.	11	14	20	24	28	9	7	21	28	24

*The two control groups are considered as one for these averages.

TABLE III

Differences Between Averages of Practice and Control Group at the Start of Experiment (Results Stated in Terms of Superiority of Practiced Group to the Control Group)

Test 1	Test 2	Test 3	Test 4	Test 5
−5 (9%)	0	+2 (2%)	+1 (1%)	−2 (1%)

Table III shows that the differences between the practiced and control groups in the preliminary trials are, with the possible exception of the results in Test 1, so slight that the groups may be considered of equal abilities in the tests at the start and therefore comparable groups. The fairest method of comparing the two groups at the end of the experiment is to figure, in each case, from

the average of the performance of all individuals (*i. e.*, in both groups) in these preliminary tests. That is, we may consider that the slight differences between the groups at the start are chiefly due to chance errors of selection and in methods of giving tests, and that the average of the records of all individuals tested is nearer to the true average than the averages of either of the two groups. The results are given in Table IV A and B.

TABLE IV A

Average Records of All Individuals (i. e., Both Practice and Control Groups) in Preliminary Tests

	Test 1	Test 2	Test 3	Test 4	Test 5
Av.	59	101	118	107	142
A. D.	11	12	24	24	29

TABLE IV B

Gains of (I) Practice Group, and (II) Control Group, in Final Tests (Compared With the Common Average of Both Groups in Preliminary Trials as Given Under A)

	Test 1	Test 2	Test 3	Test 4	Test 5
I	10 (17%)	12 (12%)	34 (30%)	14 (12%)	40 (28%)
II	10 (17%)	7 (7%)	25 (21%)	17 (14%)	18 (13%)
I–II	0	5%	9%	−2%	15%

The initial and highest records made by the practice group in the practice exercise are given in the last two columns of Table I. The practice exercise, as above described, consisted in translating sentences into code and was continued for approximately two weeks. The figures represent the number of code symbols written in the practice periods of five minutes. The average number at the start was 43, and the average of the highest records made was 239, a percentage of 537%. Although this improvement in the facility and speed with which the code could be written was considerable, the effect of this newly acquired skill on similar accomplishments was, if we may judge from the results of the control group which has no such practice, insignificant in at least certain of the tests. The results cannot, however, be taken altogether at their face value and need interpretation.

The comparisons in question are made in Table IV B. The control group in the second trial (after an interval corresponding in length to that in which the other group was engaged in practice)

did quite as well as the practiced group in one of the tests (Test 1) and even slightly better in another test (Test 4). In the three remaining tests the practiced group excelled the unpracticed or control group to the extent of 5, 9, and 15 per cent, respectively (tests 2, 3, and 5). These percentages are not, however, directly comparable. In Test 2, the percentage represents an increase in the number of operations made; in Tests 3 and 5 they are percentages of the decrease in the time necessary to do the tests.

The three tests, 1, 2,, and 4 in which the amount of improvement is computed in terms of the number of operations performed in a given time, may be considered first. This method of computation is the same as that followed in the case of reckoning the improvement in the practice exercise, the single operation,—in the first test substituting a symbol for a digit, in the second test the placing of a set of dots in a square, in the fourth test substituting a digit for a letter, and in the practice exercise substituting a code sign for a letter,—is made the unit of measurement. The question at issue is, then: Are these units equal? If they are not equal, the percentages of improvement reckoned on the basis of them are not comparable.

Specifically, the question is: Is 537% of improvement in substituting code symbols for letters equivalent, or approximately equivalent, to 537% of improvement in substituting numbers for letters?

If it is, it is at once possible to draw valid conclusions in regard to the question of the "transferability" of this improvement to other tests, *e. g.*, if there is total "transfer" we should expect to find an improvement of 537% in the digit letter substitution test as a result of the practice. Or if there is only partial "transfer" that would be indicated by the corresponding percentage.

It is to be remarked, in the first place, that there is no question but that each of these so-called units varies in the course of the experiment. It is, for example, much easier to substitute a digit for a letter towards the end of the practice than it is at the start, but both operations count "one" in reckoning the practice gain. It would, of course, be advantageous if we were dealing with "units" which did not vary, but even that is not essential for our comparison. It is simply necessary that the units in the two or more performances which we are comparing are *on the average* equal, or that we can equate them by, *e. g.*, determining how much larger or smaller one of these units is on the average than the other.

This, we think, can be done by noting the relative amounts of improvement that comparable groups of subjects can make in the same time and under the same experimental conditions in each of the tests or practice exercises. It would, of course, be preferable if the two or more groups could be practiced until the physiological limit was reached in each case—that is, until a point is reached beyond which prolonged practice does not lead, *on the average*, to any further progress. Such a level would be similar to that reached by expert stenographers or newspaper telegraphers. When one attempts to make comparisons before this level is attained he can never be sure that in one group the last few practices may have, *e. g.*, included a "sudden gain" period which has not yet been reached in the other practice. However, fairly accurate comparisons can be made on the basis of the shorter periods of practice.

In the present experiment this matter has been controlled in the case of but one of the tests. A similar class of students in the year preceding that of the experiment now being reported used the digit letter substitution test as practice material. They practiced for the same length of time as in the present experiment, and in five minute periods. A point was then reached at which the *average* improvement or fluctuations from day to day were not appreciable, and a similar point was also reached by the group in the above described practice with the code. The average percentage of improvement up to this point in the digit letter substitution is in the neighborhood of 270%.[5] It is now possible to argue that, if there is such a thing as total "transfer" of skill from one accomplishment to another, the approximately 500% of improvement made in the code practice should be evidenced by a gain of approximately 270% in the last trial of the number-letter substitution test. As a matter of fact the actual per cent. of improvement recorded in this test is 12%, which is actually two per cent. less than the gain made by the unpracticed or control group. In other words, the effect of the practice in the code in the ability to substitute digits for letters would appear, as far as these figures are concerned, to be practically nil.

The figures do not, however, tell the whole story. The introspections and observations of the subjects show that the one operation has at least interfered with the other, as observed in the tendency to

[5] This percentage is based on the records of twenty-two cases.

substitute a code symbol in place of a digit in the latter test. As a matter of fact there is doubtless an interplay of facilitation and interference, which the above used term "transfer" utterly fails to describe. The expression, it may be remarked in passing, is, in general, not a fortunate one; it has a faculty psychology connotation, and, as in this case, does not serve to differentiate between what are unquestionably very different processes. For example, the speed and skill of manipulation and the general conformity to the conditions of the experiment acquired in the code practice may very well have facilitated the performance of the digit-letter substitution test; whereas, other factors such as the specific forms of association just mentioned, to wit, digit-letter and symbol-letter substitution, may have interfered with each other to such an extent that an improvement which might have resulted from the first mentioned factors has been counterbalanced by the interferences set up by the second set of factors. In the present case, in view of the performance of the control group, it might well be argued that the interferences have in fact somewhat overbalanced the factors which otherwise might have facilitated the performance of the test. The total result, as measured by the work accomplished, remains, however, nil.

The same explanation may be made of the findings in digit-symbol substitution test (Test No. 1). In the complex dotting test (Test No. 2), there is a somewhat larger percentage of improvement made by the practice than by the control group (5%). In this case, there would appear to be relatively few factors making for interference, and the gain is doubtless due to the increased speed of manipulation acquired in the code practice. This latter factor is an illustration of the so-called "identity of substance." This term is also a rather unsatisfactory one. It is a somwehat imposing expression to describe so simple a fact, and savors, as does the word "transfer," of a faculty psychology.

In the complex dotting test, as in the digit-symbol test, the general limits of possible improvement in a single period of practice comparable to that of the code practice have not been determined, but it is apparent from the similar nature of all the tests at present under discussion that but a small part of the possible improvement has been realized as a result or effect of the code practice.

In Tests 3 and 5, both code tests, there appears to be a more considerable effect produced by the practice exercise in writing code symbols. This might have been anticipated in view of the com-

mon factors involved. What fractions of the total possible effect are represented by these percentages, it is hard to estimate. A certain fraction of the time must be used in writing down the translation, which would remain as a practically reducible minimum, whatever the improvement which might be made in the speed of translating the code. In accurate statements as regards the comparative amounts of facilitation or interference, the same procedure would be necessary as described in the case of the digit-letter substitution test.

The above comparisons have for convenience of discussion been carried out in terms of percentages. As the percentages of improvement are very evidently affected by the level at which the practice begins (*e. g.*, $50-100=$ a gain of $50=100\%$, $100-150=$ a gain of $50=50\%$), it is a mooted question as to whether or not it is preferable to make the comparisons in the absolute amounts of gain. It may be preferable in equating the units to compare the average amounts of actual gain in the two practices. In the above case, where the figures are available, the difference between the averages of the initial and highest records in the code practice is $(239-43=)$ 196, and the difference in the case of the records of twenty-two individuals in the digit-letter practice is $(291-108=)$ 183. The averages of the individual gains in the two cases are 196 and 181, respectively. On this basis of comparison, the units used to measure progress in the two practices are on the average approximately equal. The general conclusions remain unaffected by this method of comparison.

The above considerations should make evident how meaningless are the conclusions of investigations on this subject which have employed the percentages or other methods of comparison for stating the findings without proposing any method for equating the measures used.

The following coefficients of correlation have been calculated from Tables I and II.

CORRELATION COEFFICIENTS BY SPEARMAN'S "FOOTRULE" FORMULA

TEST	I	II	III	IV	V
(1) Rank Correlation between percentage gain in code practice and tests.	16±7	21±7	7±7	4±7	35±7
(a) Rank Correlation between absolute gains in code practice and absolute gains in tests	15±7	3±7	4±7	24±7	−11±7

(2) Rank Correlation between preliminary and final tests					
(A) Practiced Group	44±7	66±7	39±7	65±7	24±7
(B) Control Group	50±8	62±8	51±8	54±8	36±8

Average, Practiced Group 47.5±7; Control Group 49.6±8

(3) Rank Correlations between					
(a) Initial Record of practice in code vs. preliminary records in tests	23±7	9±7	21±7	13±7	37±7
(b) High Record of practice in code vs. final records in tests	14±7	5±7	5±7	32±7	−3±7
(c) Initial Record of practice in code vs. final record in test 5					+2±7
(4) Rank Correlation between initial and final record in practice					32±7

The correlations under (I) show that there is, in general, but little relation between the individual percentages of gain made in the practice and those made in the tests. The largest correlation is with the translation from code tests. The largest correlation is between the gains in the practice and those in the code test.

The correlations under (2) show that there is a fair amount of relation between the standings of the subjects in the preliminary and final trials of the tests, averaging in the case of the control group 49.6%, and in the practiced group a little less, 47.5%. There is comparatively little difference between the two groups on the whole, although there is some irregularity in the separate tests.

The correlations under 3a and b show what relation existed at the start between the standing of individuals in the practice exercise and in the test series and what at the end of the practice. Under both 2 and 3, a number of comparisons may be made, although the findings are not altogether consistent. The number of cases is doubtless too small, and the chance errors, in general, too great to make these detailed comparisons of value, except for purposes of illustration of methods. The results in the case of the fifth test are

worth calling attention to, however. It is quite likely that not enough time was allowed in this test to bring out the relations between the process of translating from the code and that of translating into code, and that, therefore, the coefficients of correlation found are not altogether reliable. They would appear to point to a considerable difference in the abilities of students in the two performances. The difference may be similar to that which exists between translating a foreign language into one's mother tongue and composing in a foreign language. There may very well be considerable differences in the relative abilities of students in these two processes. The subject is worthy of further study. As far as the results here go, they show (1)(see average in Table I and correlations under 1) that practice in writing code has improved to some extent the general ability of the group in the reverse process of translating from the code into English. That (2) there was less relation between the standings of students in the preliminary and final trials of this test than of any of the other tests. (That this statement holds true of the control group as well ($R = 33$) would simply argue that the two passages used for translation were not of comparable difficulty,—the discrepancy is, however, greater in the practiced group ($R = 24$). That (3) while there is a fair amount of relation (37 ± 7) between the initial record of the practice test in writing code and the preliminary test in the translation of code, there is at the end of the experiment a slight inverse relation ($R = -3, \pm 7$). One other fact needs to be considered in this connection, namely, that the rank correlation between the initial and highest records made in the practice series was $R = 32, \pm 7$. These results as a whole, in view of the various possible sources of error, can hardly be used except to illustrate the problem and the methods of attacks.

In the accompanying plates I to IV, certain characteristic curves of learning are reproduced. These have been selected from the graphs plotted by each individual from his trials in the practice series. Taking the graphs as a whole, they may be roughly classified on the basis of the rapidity of the initial progress into three groups: (1) rapid initial rise, with slower subsequent progress, *i. e.*, smaller gains towards end of practice (convex to base line), Plate I; (2) slow initial progress followed by later acceleration (concave to base line), Plate II; (3) rectilinear, *i. e.*, even acceleration from start to finish, Plate III. In Plate IV there are reproduced a number of

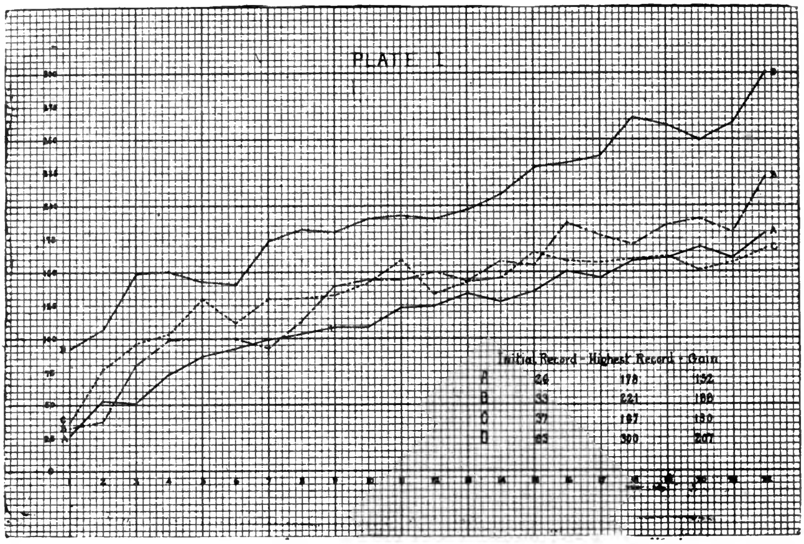


PLATE I
Initial Record - Highest Record - Gain
A 26 178 152
B 33 221 188
C 37 167 130
D 93 300 207

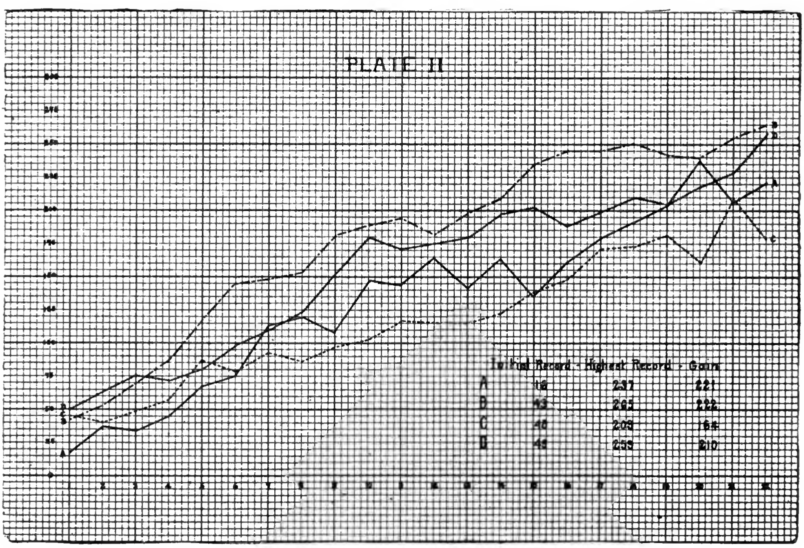


PLATE II
Initial Record - Highest Record - Gain
A 16 237 221
B 43 265 222
C 45 209 164
D 45 255 210

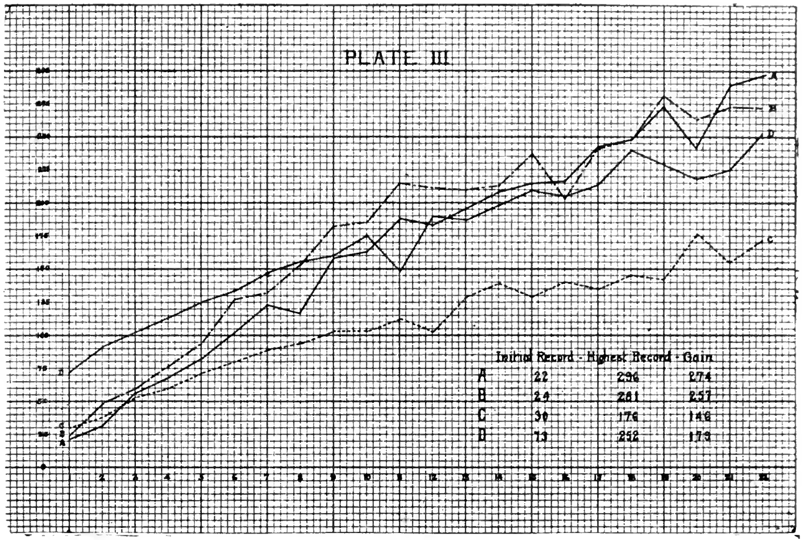


PLATE III
Initial Record - Highest Record - Gain
A 22 296 274
B 24 281 257
C 30 176 146
D 73 252 179

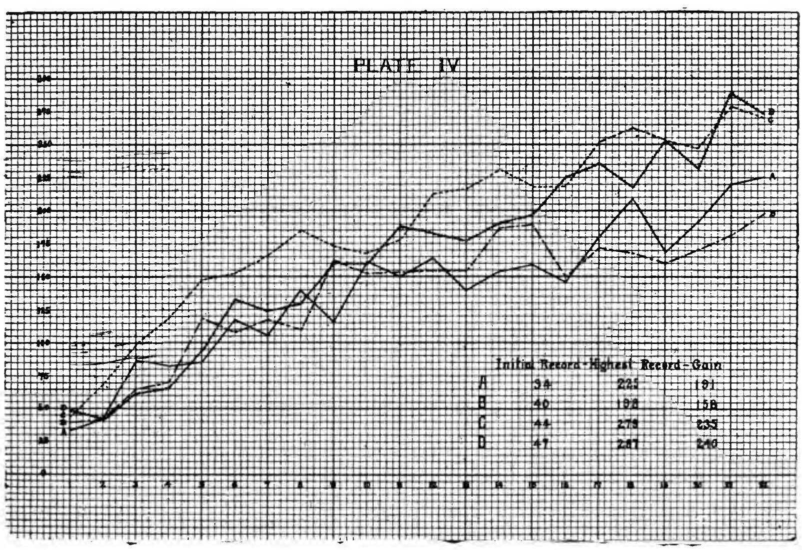


PLATE IV
Initial Record - Highest Record - Gain
A 34 225 191
B 40 198 158
C 44 279 235
D 47 287 240

curves, which show various peculiarities, *e. g.*, frequent fluctuations or "ups" and "downs" in the course of the practice, etc. The prevailing type in this code practice approximates the rectilinear with a tendency to a small initial acceleration. The initial gains are not so marked in this form of learning as in the digit-letter substitution practice discussed in the previous article, referred to above. This was to be expected, since the initial associations to be formed between the characters of the code and the letters and words of the passages transcribed are much more difficult than in the case of the associations of numbers with letters. Plateaus or prolonged periods in which there is little or no apparent gain are very common, and on the whole, of greater duration than in the digit-letter substitution practice. They are usually not followed by the abrupt or sudden gains of large amount which are often noted in the later practice curves. This again illustrates differences in the kind of learning or the sort of associations formed. In the digit-letter substitution, it is easier to acquire the digit combinations for whole words, especially for the most common words, than in the present practice. A few of these combinations, when acquired, enable the learner to leave his former level of practice as if with a leap. It is altogether likely that were this practice in the code carried out for a longer period than was done in the class experiment that the same phase of the learning would be reached more often and with the acquirement of the code combinations for whole words, the rate of progress would be capable of more rapid advance at certain stages of the learning.

Summary

Some of the general principles which may be illustrated by the experiment are the following:

(1) *Individual Differences.* If the class is of sufficient size, the results obtained may be plotted to show the range of individual differences and the effect of practice on these differences.

(2) *Characteristics of the Curve of Learning.* The results of the series of practices, when plotted, will show, in the curves of many individuals, some of the common characteristics of the learning curve, *e. g.*, the rapid initial rise, the daily fluctuations, the plateaus, and the general form of the curves, as concave, rectilinear, and convex.

(3) *Effect of Various Factors on Individual Progress.* The effect of interest, of distractions, fatigue, and rest, *e. g.*, of the Sunday in-

terval, may be studied in the curves of various individuals. The optimal length of the work period, the optimal length of the interval between the work periods and the effect of the time of day, *e. g.*, morning, afternoon, or night, on the scores, may be illustrated by dividing the class into sections and varying these factors.

(4) *Correlation.* Where several kinds of tests are used, class scores may be tabulated and various correlations computed. In certain cases evidences of negative correlation or interference may appear.

(5) *Characteristics of the Curve of Re-learning.* After the practice series has been completed and an interval of time has elapsed, a relearning series may be conducted and the relearning curve compared with the curve of original learning.

(6) *Facilitation or Interference as a Result of Practice.* In this experiment, the problem was specifically to determine whether the practice series had any effect upon the ability to the students in five other special tests. An unpracticed or control group is needed in order to test this question.

Conclusions

It is, of course, not expected that a class experiment of this nature will prove, or disprove, educational theories which are still in dispute and in regard to which experimentation is still incomplete. The experiment does, however, illustrate the problem in a concrete way and produce in the matters of individual differences, correlations, and transfer, results which are of interest.

It appears from the figures we have given that there is a positive correlation of small amount between the first tests and the final tests of students in the practice series. In other words, the students tend to hold the same relative rank in the first trials as in the last trials of the practice. This agrees with the usual findings that to those who have shall be given, *i. e.*, that those who excel at the start retain after practice their superior position. The inverse holds for the inferior students, at least to the extent that they remain inferior in rank.

In general, the correlations found between gain in the practice exercise and gain in the tests are positive, although small. The negative correlation in the case of the fifth test seems to argue for some interference between the facility gained in translating English passages into the code and the reverse process of translating the code into English.

Finally, the conclusions to be drawn from the comparative gains in the test and practice series agree in general with those of more extensive experiments on the general effect of practice.

In class experiments of this sort, it is necessary that materials be used in which the student is not already practiced and in which it is possible to improve within a relatively short practice. It would be desirable if the particular exercises chosen for the classroom tests and for the practice series represented tasks less artificial and more useful than those discussed in this experiment, and it is possible to plan such tests. Tests may be arranged in the use of the higher tables of multiplication, in using a table of squares or square roots, in learning the short methods of multiplication, in learning the use of stenographic characters, or in correcting faulty proof by means of a key of proofreader's marks. Some of these suggestions may be tried out with subsequent classes.

CERTAIN ABILITIES FUNDAMENTAL TO THE STUDY OF GEOMETRY[1]

J. H. MINNICK

University of Pennsylvania

Success in the formal demonstration of a theorem of geometry depends upon, at least, four abilities, namely:

1. The ability to draw a figure for the theorem.
2. The ability to state correctly and accurately the hypothesis and conclusion of the theorem.
3. The ability to recall additional facts about a figure when one or more facts are given, and
4. The ability to select from the available facts those that are necessary for a proof, and to arrange them so as to arrive at the desired conclusion.

The purpose of this investigation is three-fold:

1. To determine the relation of each of these four abilities to teachers' marks. This should, in turn, determine either the extent to which teachers value these abilities, or the degree to which they are able to base their marks on the things which they do value.
2. To determine the extent to which these abilities are developed in the various schools included in this investigation.
3. To develop tests which may be used for the purpose of diagnosis; that is, for the purpose of determining whether or not the weakness of a class is due to the lack of development of one or more of these abilities.

For these purposes four tests, (to be known as A, B, C, and D), have been developed, each testing one of the abilities in question. An effort has been made so to arrange each test that it will measure only one of the four abilities. Thus, Test A consisted of geometrical exercises for which the pupil was asked to draw the figure. In Test B the exercise was stated, and the figure drawn. The pupil was then asked to state the hypothesis and conclusion. In each exercise of Test C a figure was drawn, and certain facts about it were given, and the pupil was to add as many more facts as he could. The purpose of Test D was to measure the pupil's ability to select and organize facts so as to produce a proof. In order to eliminate

[1] A complete report of this investigation will be issued in book form later.

the factors involved in the other three tests the figure was drawn, the hypothesis and conclusion were started, and a list of facts, including those necessary to the proof, was given. The pupil was then asked to give the proof.

This investigation was limited to classes which had recently completed the first two books of geometry. If more advanced classes had been included, it would have been necessary to limit the investigation to these advanced classes, or to use a separate series of tests. The time required in developing a second series of tests made the latter method of procedure impracticable. To limit such an investigation to the advanced classes is to deal with a specially selected group of pupils since extensive elmination had taken place in the earlier work. On the other hand, if the abilities in question are of value, it seems that they should be fairly well developed during the first two books of geometry Also if the tests are to be used for the purpose of diagnosis it should be possible to apply them fairly early in the study of the subject, while there is yet time to remedy any difficulties which may be found to exist. Hence the tests were given to only those classes which had recently completed the first two books of geometry.

Each of the four tests was given to not less than one thousand pupils. In all, the tests were given in sixty-three schools, distributed throughout the East and the central West. The wide distribution of schools and the variation in the nature of the schools insured the inclusion of nearly every type of high school pupil in the investigation. Clearly it was impossible for the author to conduct the tests in person. However, the nature of the tests made it possible to formulate simple and definite directions which insured a fair degree of uniformity. In order to be sure that the pupils would understand these directions, the author tried them out with high school classes, and made any necessary changes before placing them in the hands of teachers, In only two cases did the returns from any of the schools indicate a variation from the directions, and the papers from these schools were rejected.

With the large number of schools so widely distributed it was impossible to prevent some variation in physical conditions, such as weather, noise, etc. However the teachers were asked to report any disturbing elements which might influence the results of the tests. From these reports it would seem that such variations were not serious.

Two scores were kept for each test. One was based on the number of necessary statements (or parts of figures in the case of Test A), given correctly, and the other was based on the number of incorrect and unnecessary statements given. The former will be referred to as the *positive* score, and the latter as the *negative* score. The author knows of no valid method of combining these two elements and certainly for the purpose of diagnosis it is desirable that they should not be combined. Weighted positive values have been assigned to the exercises of each test. These values were based upon the average per cent. of necessary statements given correctly for each exercise by all pupils taking the test, and they were so selected that a perfect score for a test would be one hundred. It did not seem possible to assign weighted negative values to the various exercises, as there was no upper limit to the number of errors that could be made, and there was, therefore, no basis for comparison. Hence the total number of unnecessary and incorrect statements in a paper was taken as the negative score.

In order to insure uniformity of scoring all papers were marked by the author. In each test the first exercise was scored for the entire school, then the second, and so on, until all the exercises had been marked for a given school. Where an exercise admitted of a variety of solutions, a copy of each solution found among the papers, together with a memorandum of how it was scored, was kept as a guide in all similar cases.

In order to detemine the relation beween the teachers' marks and the abilities in question, the coefficient of correlation between the school and test grades was computed. For this purpose it was impossible to combine the data from the various schools as they had different standards for grading. Hence the coefficient was computed separately for each school. In order to obtain this coefficient, it was necessary to array the pupils of a school according to the grades which they received for the first two books of geometry and again according to the test scores. The latter array of pupils had to be based on both the positive and negative scores. In order to meet this condition the pupils were arrayed according to the positive scores, and again according to the negative scores. The two ranks thus assigned to each pupil were added, and the pupils were again arrayed according to these sums. The formula :[2]

[2] See William Brown, *The Essentials of Mental Measurement*, pp. 42–53.

$$\rho = 1 - \frac{6 \Sigma d^2}{N(N^2 - 1)}$$

was then used to compute the coefficient and the result obtained by this formula was corrected by the formula

$$r = 2 \sin \left(\frac{\pi}{6} \rho \right)$$

TABLE I

Coefficients of Correlation for Test A.

School	r	P. E.	Relation of r to 3 P. E.
XXIII	0.313	0.068	r>3 P. E.
XXV	0.395	0.071	r>3 P. E.
XXXIII	0.111	0.180	r<3 P. E.
XXXV	0.628	0.089	r>3 P. E.
XXXVI	0.303	0.125	r<3 P. E.
XXXIX	0.426	0.140	r>3 P. E.
XL	0.487	0.079	r>3 P. E.
XLI	0.303	0.093	r>3 P. E.
XLII	0.697	0.051	r>3 P. E.
XLIII	0.436	0.088	r>3 P. E.
XLIV	0.364	0.085	r>3 P. E.
XLV	0.364	0.086	r>3 P. E.
L	0.588	0.123	r>3 P. E.
LI	0.240	0.188	r<3 P. E.
LII	0.528	0.074	r>3 P. E.
LIII	0.588	0.055	r>3 P. E.
LIV	0.150	0.147	r<3 P. E.
LVI	0.199	0.085	r<3 P. E.
LIX	0.688	0.069	r>3 P. E.
LX	0.578	0.087	r>3 P. E.
LXI	0.548	0.099	r>3 P. E.
LXII	0.436	0.053	r>3 P. E.
LXIII	0.292	0.069	r>3 P. E.

TABLE II

Coefficients of Correlation for Test B.

School	r	P. E.	Relation of r to 3 P. E.
XIV	0.467	0.127	r>3 P. E.
XV	0.548	0.127	r>3 P. E.
XVI	0.447	0.079	r>3 P. E.
XVII	0.140	0.115	r<3 P. E.
XVIII	0.188	0.056	r>3 P. E.
XIX	0.230	0.076	r>3 P. E.
XX	0.219	0.051	r>3 P. E.
XXI	0.323	0.075	r>3 P. E.
XXV	0.568	0.059	r>3 P. E.
XXX	0.395	0.081	r>3 P. E.
XXXI	0.271	0.090	r>3 P. E.
XLVII	0.668	0.065	r>3 P. E.
XLVIII	0.261	0.096	r<3 P. E.
XLIX	0.538	0.054	r>3 P. E.
LVII	0.588	0.075	r>3 P. E.
LVIII	0.344	0.086	r>3 P. E.

TABLE III

Coefficients of Correlation for Test C.

School	r	P. E.	Relation of r to 3 P. E.
VII	0.209	0.123	r<3 P. E.
VIII	0.548	0.030	r>3 P. E.
IX	0.178	0.080	r<3 P. E.
X	0.385	0.108	r>3 P. E.
XI	0.416	0.110	r>3 P. E.
XII	0.538	0.054	r>3 P. E.
XIII	0.333	0.090	r>3 P. E.
XXVIII	0.436	0.093	r>3 P. E.
XXXII	0.042	0.138	r<3 P. E.
XXXIV	0.487	0.058	r>3 P. E.
XXXVII	0.406	0.116	r>3 P. E.
XXXVIII	0.508	0.099	r>3 P. E.

TABLE IV

Coefficients of Correlation for Test D

School	r	P. E.	Relation of r to 3 P. E.
V	0.351	0.066	r>3 P. E.
XXII	0.528	0.077	r>3 P. E.
XXIII	0.209	0.078	r<3 P. E.
XXIV	0.325	0.077	r>3 P. E.
XXV	0.395	0.013	r>3 P. E.
XXVI	0.303	0.035	r>3 P. E.
XXVII	0.126	0.057	r<3 P. E.
XXIX	0.568	0.056	r>3 P. E.
XLVI	0.323	0.044	r>3 P. E.

The coefficients of correlation for each of the tests are given in Tables I–IV. If the coefficient is greater than three times the probable error it has scientific significance. In the majority of the schools, this is true, and we may conclude that usually the teachers' marks do bear a positive relation to the abilities with which the investigation is concerned. In considering these coefficients it is to be remembered that the school grades should depend upon factors not included in this investigation, and therefore the correlation should not be perfect. In several cases the coefficients are fully as high as can be expected. However they are frequently low and in many cases they are so low that, if they can be taken as indices of the value of abilities, then the schools are scarcely justified in giving as much time to this phase of geometry as they now do.

TABLE V

Summary of Positive Scores

Test	Medians for all Pupils tested	25-per-centile	75-per-centile	Lowest Median by any School	Highest Median by any School
A	62.5	51.3	72.9	50.5	78.7
B	69.3	51.8	82.2	38.5	80.9
C	50.6	36.5	65.2	29.0	67.0
D	73.3	55.4	87.0	54.7	80.5

TABLE VI

Summary of Negative Scores

Test	Medians for all Pupils tested	25-per-centile	75-per-centile	Lowest* Median by any School	Highest Median by any School
A	7.1	10.7	4.2	11.8	4.8
B	3.5	5.9	1.5	4.5	2.0
C	4.1	7.3	1.9	7.3	2.4
D	2.6	5.4	0.9	3.7	1.5

The median positive and negative scores made by all pupils taking each test are given in Table V and VI. Since we do not know that the tests measure the respective abilities in the same way, we can not compare the results of the different tests. Thus the low median score of Test C does not necessarily indicate that the pupils have less ability to recall facts about a figure than to do any of the things called for in the other tests. However, we may compare the results obtained by giving the same test in different schools. Such a comparison shows that some of the schools are developing these abilities much more effectively than others. In the case of each test the marks of some schools are quite satisfactory, while those of others are very low. This variation in achievement may be due, in part, to differences in local conditions, rather than differences in methods and in teaching ability. Nevertheless, it is difficult to see how local conditions could result in the extremely low scores made by some schools. If the abilities tested are essential to success in the study of geometry, then the results indicate that progress is almost impossible in some of the schools until these abilities have been further developed. On the other hand, the achievements of other schools indicate that it is altogether possible to develop these abilities to a fair degree during the study of the first two books of geometry.

*As the negative scores represent the numbers of incorrect and unnecessary statements, the larger numbers represent the lower scores.

Since, as we have seen, these abilities may be satisfactorily developed, but in many cases they are not, then if they are of value it is important that we have some means for determining when they are satisfactorily developed. It is believed that these tests or others similar to them will prove valuable for this purpose. The standards given in Tables V and VI furnish a means of comparison. If a class is above the median scores in these tables, and especially if it is near or above the 75-percentile marks, the teacher may be fairly sure that the abilities have been sufficiently developed to insure the success of the class. If, however, the class falls below the median score in any one of the tests, and especially if it falls near or below the 25-percentile mark, special effort should be made to develop the ability in question.

A STUDY OF A CLASS OF CHILDREN OF SUPERIOR INTELLIGENCE

HENRIETTA V. RACE

Psychological Laboratory, Louisville Public Schools, Louisville, Ky.

The idea of special classes for children with traits markedly different from the average child has been advocated by advanced thinkers in education for several years. Its first realization was in the organization of classes or schools particularly suited to deaf, dumb, and blind children. Later, notably in the first years of the twentieth century special classes for the feeble-minded were formed. During the last decade and especially during the last five years, attention has begun to turn to the children at the opposite end of the scale from the defective, and those children who show unusual ability are coming into the field of educational research and special pedagogical endeavor. Dr. Edward L. Thorndike, in an address in New York in 1914, stated that the United States government and numerous organizations of the state and city, to say nothing of private societies, were spending millions of dollars, doubtless justifiably, for the care of those who were low on the scale of mental ability, such as criminals, paupers, and defectives in mind and body, but that nothing was being done in state or nation in the way of special care for those who were unusually gifted. Dr. Thorndike stated further that, in his opinion the benefit which would accrue to the state in the effort to preserve and to develop to the extent of its ability its best product would far out-value that derived from the care of the less able. Since that time, students of mental development, especially those working in educational psychology, have shown a keen interest in the gifted child and his opportunities. Here and there special effort has been directed to the training of gifted individuals, but to the writer's knowledge, no work had been done in the public schools for the gifted child, comparable with the specialized effort made to train the feeble-minded, until within the last year.

In October, 1916, a class for exceptional children was organized under the control of the administration of the city schools of Louisville. For two years Louisville had had an accelerated class where children of the fourth grade, who in the judgment of teacher and principal were considered especially able, were allowed to complete in one-half year the work prescribed by the course of study for one year. It was found, on investigation, that these children differed

greatly in mental caliber. Some were over age; some normal; while others were under age. Not all were exceptionally capable.

In order to get a class, each member of which had unusual natural ability, it was decided that the organization must be made on the basis of scientific mental tests. Consequently, during the month of September sixty-two children were examined; fifteen were selected who had an intelligence quotient ranging from 120 to 145. It was thought that the number was too small for a separate room and teacher, so they were taught with the accelerated class until February. These children accomplished one year's work of the curriculum in a half year and were the leaders in the class room exercises. In February one hundred eleven more children were examined. Twenty-two were selected and the "Opportunity Class" at the Normal School was organized. This group consisted of ten boys and eleven girls. They ranged in age on entrance in February from 7.7 to 9.8 with an average age of 8.7 years. Almost all of them were just beginning the work of the 4B grade though three, numbers 1, 8 and 19, were entering 3A. The Terman Revision of the Binet-Simon Scale test supplemented by the Goddard form board was made the basis for entrance. Table 1 shows the result of the measurement in intelligence quotients.

TABLE I

No.	Boys	No.	Girls
1	159.1	11	136.6
2	145.4	12	141.6
3	134.0	13	135.2
4		14	145.6
5	142.3	15	161.9
6	137.3	16	120.5
7	159.4	17	122.3
8	155.1	18	137.4
9	167.9	19	135.8
10	138.0	20	123.4
		21	123.0

The range of ability is from 120 to 167.9 I. Q. with a median of 137.4, standard deviation of 13.9 and probable error 8.1. Dr. Terman finds standing of from 90 to 110 I. Q. to be normal, from 110 to 120 to be superior intelligence, from 120 to 130 very superior intelligence and from 130 on genius or near genius.

During March, April and May other tests that were given were the form board standardized by Dr. Goddard, the analogies test, by Dr. Whipple, the easy and hard direction test, the genius-species

TABLE II

	M		S. D.	P. E.
Analogies B				
Gifted Class				
March	18.9	Words	1.4	1.4
	M– 3.5	Sec	1.5	.7
	Av– 7.0	Sec		
May	20.0	Words	1.0	.0
	M– 2.7	Sec	1.0	.6
	Av– 4.2	Sec		
Analogies B				
Whipple, Selected Children	Av– 8.9	Sec		
Unselected "	Av–11.11	Sec		
Opposites				
Gifted Class				
IV	19.5	Words	1.1	1.4
	7.0	Sec	1.8	1.5
IV Op	20	Words	.07	.0
	6	Sec	1.4	.8
Opposites				
Normal (Norsworthy)				
9 years				
IV	9.5	Words	....	1.7
IV Op	9.0	Words	....	2.0
Genus Species	20	Words	2.3	0.0
Gifted Class	8.	Sec	3.3	2.2
Genus Species				
9 years				
Pyle	10.8	Sec	2.9 A. D.	
Easy Directions	Perfect Reaction		....	....
Gifted Class	6.9	Sec	1.2	.9
Woodworth & Wells				
College Students	3.6		.92 A. D.	.28
Hard Directions				
Gifted Class	18.5 reactions		2.1	1.8
	10.6 sec.		2.4	2.0
Hard Directions				
Woodworth & Wells	5.4 sec.			
College Students				

Form Board			
Gifted Class	18.1	2.2	1.5
Form Board			
9 years	18.2	....	....
Goddard			

M—Median. S. D.—Standard Deviation. P. E.—Probable Error.

test by Woodsworth and Wells, and the opposites tests used by Dr. Norsworthy. Table II shows the results in quantity of work and time, in comparison with available records for the average child. The children were told, except in the analogies test, to write as rapidly as possible the requirements of the test. The record is the median word accomplishment and the median of the average association time. The analogies test was given orally and the reaction time for each element was taken.

Table III shows the correlations of the above results by Pierson coefficient.

TABLE III

I. Q. and Opposites	.662
I. Q. and Analogies	.57
Analogies and Opposites	.75
Analogies and Genius-Species	.68
Analogies and Hard Directions	.43
Analogies and Easy Directions	.36
Analogies and Form Board	.20
Opposites and Genus-Species	.62
Opposites and Hard Directions	.36
Opposites and Easy Directions	.62
Opposites and Form Board	.38

These tests were chosen because of the varied abilities which they seemed to measure. The results show that children of mental acumen ranging above 120 can do work at least two or three years in advance of the average child, and can perform varied kinds of mental activity almost equally well. An exception must be made of the form board tests in which ideas function least and motor control enters quite extensively.

Besides the individual measurements, tests were also made with educational scales. The results with standard comparisons are given in Table IV.

The correlation of the educational tests with the intelligence quotients is given in Table V.

TABLE IV

	M	S. D.	P. E.
Reading (Thorndike) Gifted Class	6.6	.6	.25
Thorndike Standard			
4th Grade	5.25		
6th Grade	6.5		
Reading (Kelly) Gifted Class			
Test for grades 3, 4, 5	20.4	4	1.9
Test for grades 6, 7, 8	20.	6	3.0
Kelly Standard			
Test for grades 3, 4, 5	9.9		
Test for grades 6, 7, 8	13.4		
Language Trabue, gifted class Scale B	16.	2.0	1.0
Trabue Standard			
4th Grade	3.5		
6th Grade	4.5		
Spelling (Ayres) Gifted Class	25 Words 98%	1.0	.0
Ayres Standard			
4th Grade	88%		
6th Grade	98%		
Reasoning Problems (Starch) Gifted Class	7.7	1.3	1.5
Starch Standard			
4th Grade	6.2		
6th Grade	9.4		
Addition (Courtis) Gifted Class	7.0	2.0	2.2
Courtis Standard			
4th Grade	5.		
6th Grade	9.		
Addition Woody	14.	2.7	2.7
Woody Standard			
4th Grade	11.		
6th Grade	16.		

TABLE V

Thorndike Reading and Intelligence Quotient				.494
Kelly Reading	"	"	"	.39
Trabue Language	"	"	"	.509
Starch Reasoning in Arithmetic	"	"	"	.48
Woody-Addition	"	"	"	.14
Courtis Addition	"	"	"	.09

Judging from the results of the educational tests, children of 120 I. Q. and above can do school work at least two or three years in advance of the average child and can master the different studies of the curriculum almost equally well. Reading and language contests stand at the head of the list and mathematical work as it is found in addition and reasoning problems the lowest. This latter seems to be due to the more extensive exercise of the former function and to lack of practical training in the latter.

Table VI shows the physical measurements of the children under consideration in height and weight. These measurements were taken in March. They are not accurate, since the children were not stripped, but they were given a fair estimate of physical condition and show that the members of this class exceed the average child in height and weight.

TABLE VI

Boys	Chro age	Weight Kg	Normal Wt. Kg	Height Cm	Normal Ht Cm
1	...	.....	.....	.....	......
2	8.5	28.1	25.06	134	123.48
3	8	27.2	23.81	129	120.93
4	8.5	22.0	25.06	123	123.48
5	...	....	.....	.....	......
6	8.5	34.4	25.06	141	123.48
7	8.5	28.8	25.06	133	123.48
8	...	....	.....	.....	......
9	8.5	23.6	25.06	121	123.48
10	8	27.6	23.81	124	120.93
Girls					
11	9.5	38.7	26.52	143	127.74
12	...	....	.....	.....	......
13	8.5	22.0	24.22	129	122.75

14	8	21.1	23.01	120.5	120.49
15	9	27.4	25.25	131.5	125.24
16	8.5	22.9	24.22	127	122.75
17	8.5	23.8	24.22	125	122.75
18	9	31.0	25.25	128.5	125.24
19	8.5	26.1	24.22	134.5	122.75
20	9	20.2	25.25	123	125.24
21	8.5	25.2	24.22	149	122.75
Average	8.5	26.5	24.66	130.3	123.35

In character and disposition these children are conceded by all who have worked with them to be superior. They are not conceited or puffed up by their selection or work in the class. However, improvement in this respect has been marked. The teacher was asked to take note of morale in order to ascertain if possible the value, in terms of character development, of the gifted class segregation—a child's opportunity to work with his peers. The notes are given below without editing.[1]

"I have noticed changes in the dispositions of many of the children. The following are most worthy of note.

"A constantly spoke in a surly, critical tone and did not enter into the activities of the class with apparent interest. He seemed to be somewhat on the defensive. Now he shows a co-operative spirit and we have not heard that surly, complaining tone for two months or more.

"B was exceedingly nervous and so self conscious that she stammered when called upon to recite. She is much less nervous, and speaks with assurance and self possession.

"C has always been lovable and thoughtful of others but at first, because of excessive interest he was inclined to insist upon dominating the situation during discussion. Now, he considers the rights of others and is satisfied to listen to the contributions of other pupils and to await his turn.

"D was conceited, overbearing, and hypercritical. He now shows co-operative spirit and is not at all critical."

The school work of the Opportunity Class represents in value, according to the Louisville course of study, the performance of the entire fourth year. Besides the accomplishment of this work the children learned to use with a considerable degree of freedom four hundred words in conversational German. They also composed the poem and music of a spring song and an operetta.

The class did this work happily and with ease. Home study was discouraged except where it was a matter of great desire and then it was limited to twenty minutes.

The school room was made as attractive and homelike as possible. It was seated with chairs so as to allow freedom of movement and furnished with library tables for reading. Plants and flowers filled the banked windows. The whole atmosphere of the room was that of happy, joyous work-a-day freedom.

The method of teaching employed might be termed the method of vital experience. Abstract ideas were, in no sense of the word, the stock in trade but the needs of the group in the normal life activities, real situations in the social milieu were made the foundation of the learning process.

It would be unfair to close this paper without some tribute to Miss Jessie Marshall, the young woman who served as teacher of the class and whose life closed with the school year. Her work with the class was truly a masterpiece of teaching art. She "lived with the children" in the best sense of that ideal; giving them the privilege of growth by their own experience and making the way of progress delightful and exceedingly profitable.

Conclusions: From the results of the different tests and the experiences of the year it seems justifiable to conclude:

First, that gifted children are able to accomplish, with ease, the ordinary two year's school work in one year.

Second, that they are apt to be unusually able in various fields of human learning.

Third, that they are especially capable of handling ideas and that their thinking is marked with quickness and directness.

Fourth, that gifted children in common with those of less mental calibre can and do form bad habits thereby crippling their ability. This was especially shown in the studies in arithmetic and social attitude.

Fifth, they are highly social in the scientific sense of the term. They tend to have good dispositions and lend themselves generously to the needs of the group.

Sixth, their lives should be normal and natural. Abstract ideas should not be their stock in trade but living vital experiences. They should be allowed to think for themselves.

Seventh, in order that they may contribute most acceptably, gifted children should have special opportunities to work in cooperation with the social group.

COMMUNICATIONS AND DISCUSSIONS

AN APPARATUS FOR THE MIRROR-DRAWING TEST

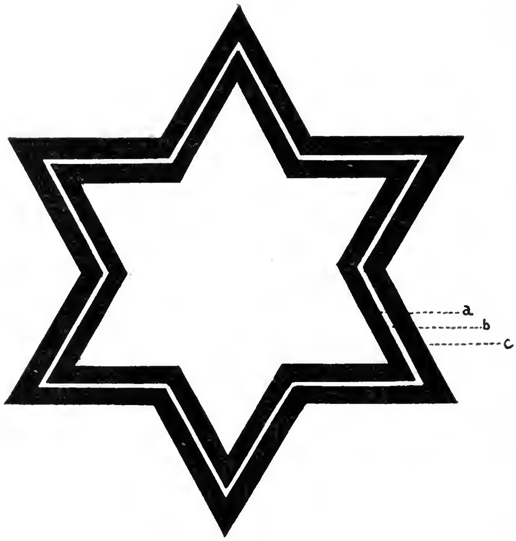
The interesting and valuable results of mirror-drawing tests are often obtained only at the cost of a good deal of labor. The usefulness of this test is much restricted by the fact that so much work is required after the test has been completed to find out how many errors have been made. And just because so much labor of computation is unavoidable in experimental psychology it is worth while to eliminate it wherever possible. The figure in Whipple's *Manual of Mental and Physical Tests*, test 36, page 493, is quite typical. An attempt to determine the number of errors in that figure is sufficient to emphasize the fact that when these errors have to be actually counted on many such figures the expenditure of time and effort required is not small. I have found that it takes really more time for students who are working upon mirror-drawing problems to compute their errors than it does to make the mirror-drawings in the first place.

To obviate this difficulty the following very simple arrangement has been made and is being successfully used in the psychological laboratory of Ohio Wesleyan University.

Two concentric stars of metal are made and so arranged that the outside line of the smaller shall be smaller than the inside line of the larger. Thus when they are arranged concentrically the space between them is itself star shaped. If this space be filled up with some plastic substance which becomes very hard and which differs in color from the metal a clear cut and permanent star will be formed. The metal strips are electrically connected with each other and to an electric counter, arranged in series with some power source, and with a stylus having a sharp metal point. Add the familiar mirror and screen.

The test is performed in exactly the same way as if the printed star outline and pencil were used. The student tested simply attempts to keep the stylus upon the inscribed figure. Any and every time the stylus moves from the inscribed star the counter registers. And at the conclusion of the test it is as easy to determine the number of errors as it is to determine the elapsed time. In each case one simply reads a dial.

A point of importance is that by means of the apparatus a somewhat simple definition of an error is made. An error consists of enough lateral movement of the stylus to move it off of the inscribed star. Thus an error means exactly the same thing to all students, and its definition will not vary from day to day.



a and c are the brass stars, five millimetres in width. b is of non-conducting material, one and a half millimetres in width. a and c are electrically connected to a binding post set in the upright (not shown) at the left rear corner of the cement base.

The apparatus as used in the psychological laboratory of Ohio Wesleyan University consists of a cement base in which are embedded two brass stars, arranged concentrically as described above. The resulting inscribed star measures ten centimetres from point to point, and is one and a half millimetres in width. It is needless to say that the cement is very carefully smoothed so that its surface is exactly level with that of the brass stars. The strips of which the brass stars are made are five millimetres in width. At the back and left of the cement base a metal post is inserted for the support of a mirror which is so pivoted that it can be tilted and fastened at any angle from the perpendicular. From this upright an arm projects to support the small metal screen which intervenes between the eyes

and hand of the one taking the test. The electric connections from the brass stars run under the base to the mirror upright, and a binding post in the upright serves to connect the stars with the batteries.

Obviously the form of figure used may be varied as desired. It is also possible that in later types some more suitable material than cement may be found for the base.

The handiness of this little apparatus may enlarge the usefulness of the test. It has already been suggested that the test is capable of being employed as a test of quickness in learning. If so used this form of apparatus will reduce the amount of work and the time required for performing the test very considerably.

G. R. Wells.

Ohio Wesleyan University.

PERCENTILE NORMS FOR SCALING DATA

Scales devised upon a basis of percentile differences are bound to be very crude, besides embodying a fundamental fallacy. Measurements of a group may be resolved into quartile, quintile, decile or percentile intervals. A subject measured for a given quality may conveniently be compared with his group to determine his location within the group. Or two individuals measured for the same trait may be compared for relative differences in the light of the percentile grades of the group. We cannot, however, equate, translate, or combine the raw percentile values in various tests, because the significance, for example, of the 80th percentile may vary greatly when we consider the variability and dispersion of the different traits.

But what is perhaps of more fundamental importance is that for testing purposes the unrefined measurements obtained for a test should be first transformed into P. E. (probable error) values, or into sigma (standard deviation) values. These only can give us absolute units necessary in scaled data. Excellent examples of this type of procedure are the studies of Buckingham, Trabue and Woody. In experimental analyses of this sort one is certain that the difference between points 5 and 6 on our scale is equal to the difference between 0 and 1, or 14 and 15, or any other two adjacent units. This, however, is not true for units in a percentile scale. The difference between the 0 and the 10th percentiles is not equal to the difference between the 50th and 60th or between the 70th or

80th percentiles. This fact will be obvious to anyone from a mere cursory examination of the ogive. Consequently the estimate of an individual's ability based on percentile differences in various tests which are finally averaged must not only yield but a very crude measure of intelligence development but is also to an extent misleading. This is emphatically true where an arbitrary score is assigned to various performances differing in the amount of error.

SAMUEL C. KOHS.

Leland Stanford Junior University.

ABSTRACTS AND REVIEWS

AUGUSTA F. BRONNER, PH.D. *The Psychology of Special Abilities and Disabilities.* Boston: Little, Brown and Company, 1917. Pp. VI, 269. $1.75

The rapidly growing structure of differential psychology which is developing about the corner-stone of individual differences has received a somewhat distinctive and valuable addition at the hands of Dr. Bronner. Too little attention, as this new work emphasizes, has been devoted to specialized powers and defects as distinct from general proficiency and general deficiency of intelligence. To the recognition of variability in individuals within the group must be added the recognition of variability in the relation to traits within the individual. Dr. Bronner with this end in view proposes to discuss examples with their significance, both of (a) particular disabilities in those who have normal general ability and of (b) particular abilities in those who have less than normal general ability, together with methods of attacking both these groups of cases.

The major emphasis of the study is throughout upon individual case examinations to which, in fact, two-thirds of the pages are devotvoted. Methods of examining the individual cases and precautions to be observed in the interpretation of results of these examinations are set forth in two chapters on diagnosis. An intimate and comprehensive study of each individual by means of a wide range of tests is to be supplemented by careful interpretation of the test records in the light of disturbing factors and in conjunction with an analysis of heredity, developmental history, and environmental conditions. The mere administration of any "intelligence scale" is never suf-

ficient for trustworthy diagnosis: for a pronouncement of feeblemindedness, the Binet results must be borne out by results with performance tests, by reactions to common sense situations, and by the extent of profiting from educational opportunities. In an interesting chapter on present educational tendencies attention is again called to the necessity for intimate study of individual cases as the prime need in analysing school and vocational failures and maladjustments.

The remainder of the book considers many rather detailed reports of examinations, grouped into types illustrative of specific defects and abilities, and sets forth suggestive ideas upon the practical social aspects of the cases. Such defects are treated as those which occur in number work, language ability, separate mental processes affecting all work (as memory, manipulation of concrete material, general slowness of reactions, higher processes), and defects in mental control. Instances of special abilities are reported occurring with general subnormality in fields of mental operation closely analagous to those in which these defects occur. In order to discover these special abilities and disabilities it is necessary, of course, first to possess a thorough analysis of the mental processes involved in the field where the defect occurs and then, second, to test for each such processes by rather specific tests. This is the plan of attack which Dr. Bronner approximates in each type of defect considered. All the forty-six cases given as examples in the text are succinctly summarized in an appendix of "Records of Psychological Examinations."

The conclusion of the entire study echoes the introduction in its emphasis upon the necessity for comprehensive diagnosis of the individual. As may well have been inferred from the nature of the treatment throughout, Dr. Bronner does not indulge in any glittering generalizations as a conclusion of her work. Perhaps in this respect the book does not quite satisfy the high hopes raised by the preface, but at least the abundance of factual material is a welcome substitution for fine theories and empty generalizations.

ARTHUR W. KORNHAUSER.

Carnegie Institute of Technology.

The Journal of Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD.

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

THE JOURNAL is issuing a series of Monograph Supplements for the publicaiion of more extensive researches in educational psychology. For details prospective contributors may apply to Warwick & York, Inc., Baltimore, Md.

The price of the JOURNAL is $3.00 per year of ten monthly issues, if paid in advance. Single issues are 50 cents each. Back volumes are $3.50 in original wrappers and $4.75 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc., 10 E. Centre St., Baltimore, Md.

VOLUME IX FEBRUARY, 1918 NUMBER 2

EDITORIAL

SCHOOL POLICIES IN NEW YORK CITY

The recent change of administration in New York City carried with it a very considerable reorganization of the school system. One of the chief issues of the campaign was the retention or the rejection of the so-called Gary plan. The overwhelming character of the vote gave the new administration a mandate to eliminate the Gary plan root and branch, and to undo all the work done in that direction in the past six years. This is not easy, and it is to be hoped that some of the good features of that scheme may be retained or brought back into the system under another name.

One of the first tasks confronting the new mayor was the appointment of a school board of seven members to take the place of the old board of forty-six members. The selection of the personnel of this board was obviously a matter of some moment, since upon it would rest the responsibility for the administration's educational

policy, and upon the success of its measures would depend the public approval or disapproval of the change from the large to the small board. Of the mayor's appointees only two had ever been on the school board before, and only one, Mr. Arthur S. Somers, subsequently chosen president of the board, had seen continuous service in recent years. Thus most of the members were new to the educational public of the city, and their attitude was a matter of no little speculation. Within a few weeks Superintendent Maxwell, who has served the city faithfully for more than a quarter of a century, and who is now in poor health, was made superintendent emeritus under generous conditions, and the board began its search for a new superintendent. The selection of the proper man for this position increases the burden of responsibility upon the new board and is fraught with momentous consequences for the schools of the city.

In view of these facts the inaugural address of President Somers, who by his experience and his position has a preponderating influence in shaping the policies of the board, is a noteworthy document. The old board in spite of its large membership was always overwhelmed with a mountain of work, and was thereby greatly delayed in the transaction of business. Many of those in close touch with the New York schools contended that a board of seven members was entirely too small, and could never do the work that necessarily devolved upon the board of education in such a large city. On this point President Somers takes a wise and statesmanlike position: "A board of education of seven members cannot undertake, in the same manner, to do the work that a board of forty-six members did. The mass of unimportant matters and detail that came before the preceding board, because a system of standing committees encouraged it, we shall not set out to consider. We shall be a policy-making body. We shall not undertake to run the schools, but we shall make it our duty to see that they are run properly. To do this effectively we must abandon standing committees. They encourage the reference to the members of too many matters of routine. They lead department heads to seek committee approval before proceeding to action and thereby make for divided responsibility. If seven members are to undertake to render the largest possible service to the city schools, they must leave to the paid officers all matters too detailed for the consideration of all the members meeting together. They must work through a committee of the whole, rather than through standing committees, but

when occasion demands, or a particular problem is to be solved, it may be necessary to call upon a smaller number, as a special committee, to study the proposals. This board of education must, as I see it, determine policies, leaving to department heads the duty of carrying them out. We must hold our paid officials, all of them, to a strict accountability for results, and we shall not undertake to do other than give them the largest measure of freedom." This is in accord with the most approved findings in educational administration, and augurs well for the New York schools.

But not only are the members of the professional staff promised freedom from interference by board members in administering the work of the schools, but they are urged to free themselves from the burden of petty detail, in order that they may more intelligently study the problems that confront them. "As this board of education will not undertake to do what the large board that preceded it tried to do, but will consider problems and policies and not details, just so our board of superintendents and our district superintendents, our board of examiners and our directors, would do well to study ways to rid themselves of the burden of considering petty matters and numerous details. It has occurred to me to suggest for the consideration of the superintendent of schools and the board of superintendents that the members of that board be assigned to particular problems rather than to certain parcelled out branches of the service. When a superintendent is assigned to a committee on elementary schools or to a committee on high schools he gradually becomes submerged in a mass of routine and detail that hampers his vision of bigger problems. Similarly, our district superintendents have become too much examiners of details of organization and too little leaders in the solution of great school problems. The tendency is to become more expert in rating teachers and classes than in improving the teaching of children. The former, to my mind, is a relic of antiquity that should be among the first customs to be discarded. I believe that much time is spent in examining and rating teachers who are fully qualified. If that be so, why should not the principal's rating be sufficient? In any case where there is a doubt or where injustice is claimed, appeal may then be made to the superintendent, but only in such cases. Far more important is it that they concern themselves with the contents of children's heads than of teachers' desks and plan books. All this is a big problem for the board of superintendents and the district superintendents to study,

and I earnestly invite such study. Possibly it would be well to test a reorganization under which district superintendents would be assigned not to districts of varying size and unequal register, but to associate superintendents directed to study particular school problems."

Let us hope that this will mean the beginning of an era of the scientific study of educational problems such as New York has never seen. The problems lie in abundance on every hand, and there are many men on the supervisory staff well able to conduct such scientific studies. What is needed is a consistent and rational policy that will encourage such studies, and will demand them as part of the routine of supervision. The recent action of the board of superintendents and the board of education in authorizing the publication of 30,000 copies of the new New York penmanship scale, so that a copy may be posted in every school-room, is a step in the right direction. But let us also have studies of oral and silent reading, fundamental operations and reasoning in arithmetic, of spelling, of geography, of history, of literature, of elementary science. Let us actually determine the things that groups of children are able to do under definite school conditions; let us grapple with the problems of the bright child and the dull child and modify our school work to meet individual needs; let us ascertain to what extent school brightness and dullness are general traits and to what extent they depend upon specific characteristics; let us, in fine, demand that our educational procedure be based upon ascertained fact rather than upon traditional belief and vague opinion. J. C. B.

NOTES AND NEWS

An interesting development in educational scales and measurements is promised in the forthcoming New York City Scale of Penmanship, which is being devised by Clyde C. Lister and Garry C. Myers of the Brooklyn Training School for Teachers. Out of several thousand samples of penmanship collected from New York City school children, three hundred were selected as providing an even range from the best to the poorest. These three hundred samples were then sorted by each of twenty-five competent judges (most of them teachers of penmanship) into eight piles, the intervals between each grade being made as nearly equal as possible. This sorting was done three different times, first from the point of view of the form of the letters, second from the point of view of the spacing of the letters and words, and third from the point of view of the movement involved in the writing. Elaborate directions were given for judging the samples from each point of view, and after each sorting the samples were recombined in chance order. In the final arrangement there will thus be three scales, one for form, one for spacing, and one for movement. Each will consist of eight specimens, so selected on the basis of statistical treatment of the judgments as to constitute equal steps in a scale of increasing excellence. A detailed account of the construction of the scale will appear in an early number of THIS JOURNAL. The scale has been approved by the New York City Board of Education, and the printing of thirty thousand copies of it has been authorized.

An important question in educational psychology is raised in the following inquiry from a correspondent. "What would be the advisability of placing permanently upon the blackboard before children of the third to the eighth grade such sentences as the following: "I saw, not I seen;" "I did, not I done;" "He and I, not him and I;" and so on? Would this be a good method of correcting poor English? What would be the effect of the constant presence of the correct form, of the incorrect form, or of both together upon (1) the child who habitually uses the incorrect form, and upon (2) the child who rarely or never uses incorrect forms?" This sets a neat task for an educational experiment. Perhaps the chief difficulty in control would be the degree to which the teacher brings the forms to the focal attention of the pupils. If the forms were merely written on the blackboard and allowed to remain there without attention being called to them, they might have no more effect than the pictures on the walls or other articles of school furniture. This, however, might be controlled by systematic attention to the forms at stated

intervals. The question is a significant one, and the JOURNAL will be glad to receive for publication any experimental data bearing on it.

Public School 166, New York City, Miss Margaret Shaw, principal, reports an "Opportunity School" where children are given a chance to develop according to their abilities. The regular school curriculum is followed, but the school hours are from 8:30 until 4, and each pupil receives at least 100 minutes per week of special coaching in a "rapid advance" coaching class. The work of the school is supplemented by that of the Children's House, which is also under the direction of Miss Shaw. The average attendance is about 350, and last year 106 pupils graduated from the school, all of whom had done six terms' work in four terms or less.

The War Department has asked the land grant and technical colleges of the country to undertake the instruction during the summer of students from the military and training camps.—*School and Society.*

A Committee on Education and Special Training has been created by the War Department to study the needs of the various branches of the service for skilled men and technicians; to determine how such needs shall be met, whether by selective draft, special training in educational institutions, or otherwise; to secure the cooperation of the educational institutions of the country; and to administer such plan of special training in schools and colleges as may be adopted. The committee consists of Colonel Hugh S. Johnson, deputy provost marshal general; Lieutenant Colonel Robert I. Rees, of the General Staff; and Major Grenville Clark, of the Adjutant General's Department. The committee will be assisted by the following civilians: Charles R. Mann, Massachusetts Institute of Technology, James R. Angell, University of Chicago, J. W. Dietz, Western Electric Company, president of the National Association of Corporation Schools, James P. Monroe, member of the Federal Board for Vocational Education, and Samuel P. Capen, United States Bureau of Education, specialist in higher education.—*Science.*

At a recent meeting of the University of Michigan board of regents Professor Whitney reported an increase in enrolment in the department of education and a decrease in the teaching force, and requested the appointment of a professor of educational psychology and experimental education.

At the recent meeting of the Education Association of Western Pennsylvania Professor G. D. Strayer, of Columbia University,

spoke on "The Place of Scientific Measurement in Education," and Professor C. H. Judd, of the University of Chicago, on "Non-promotion." The subject of "Measurements" was the topic of discussion at fourteen separate section meetings.

The University of Rochester has expanded its work in psychology. Quarters are now provided for an experimental laboratory, and are thoroughly equipped for experimental purposes. Two experimental courses are given during the present year. One course, extending through the college year, emphasizes the psychology of the sense organs and more complex mental processes. The second course takes up the study of comparative psychology. Quarters for animal experimentation have been provided. The course is under the charge of L. A. Pechstein, Ph.D. (Chicago).

Dr. J. E. Wallace Wallin, director of the Psycho-Educational Clinic and Special Schools in St. Louis, will offer courses in Clinical Child Study and Subnormal Children in the summer session of the University of California in Los Angeles.

Professor Frank M. Leavitt, of the University of Chicago, has been appointed one of four associate superintendents and director of industrial training in the schools of Pittsburgh. He will also teach in the Carnegie Institute of Technology.—*School and Society.*

Dr. C. Judson Herrick, professor of neurology in the University of Chicago, has recently been commissioned major in the sanitary corps of the National Army.

Dr. R. H. Sylvester, assistant professor of psychology, University of Iowa, has joined the psychological division of the medical officers' training camp, Fort Oglethorpe, Ga.

Assistant professor Ambrose L. Suhrie, University of Pennsylvania, has been appointed head of the school of education, Western Reserve University, Cleveland, Ohio.

Dr. Francis N. Maxfield, assistant professor of psychology and assistant director of the psychological clinic, University of Pennsylvania, has been appointed psychologist in the public school clinic of Newark, N. J.

At the Madison, South Dakota, State Normal School Professor W. J. Bell has been appointed assistant director of the training school and chairman of the bureau of tests and measurements.—*School and Society.*

PUBLICATIONS RECEIVED

CHARLES ALTSCHUL. *The American Revolution in our School Textbooks.* New York: George H. Doran Company, 1917. Pp. 168. $1.00.

The war is changing our ideas about many things. It is making us more critical about the ultimate consequences of our educational procedure. The old isolation of America from the rest of the world, so urgently preached by Washington, is gone forever, and now we are beginning to realize that through the emphasis of half truths in our school histories we have deliberately built up a prejudice against some of our best friends. It is the service of the present study to bring this fact out clearly and definitely, and to show the need for a broader and juster treatment of the American Revolution. The author has made a careful study of ninety-three school histories published between 1880 and 1916, and has divided them into five classes on the basis of the fullness and fairness with which they present the conditions of thought in England and America prior to and during the Revolution. The book is made up largely of quotations from these texts and comments thereon. One soon becomes aware of the distorted and norrowly chauvinistic view which is inculcated by such texts, and appreciates the author's demand for a treatment which will bring out the fact that there was a parallel struggle between liberal and reactionary ideas in England itself, and that the American uprising was but a phase of the general movement toward democracy which later overturned France and England, too. The book should be in the hands of every teacher of American history.

WALTER S. ATHEARN. *Religious Education and American Democracy. The Correlation of Church Schools and Public Schools. A Community System of Religious Education.* The Malden Leaflets, Numbers 1–3. Boston: The Pilgrim Press, 1917. Number 1, Pp. 27. Fifteen cents. Number 2, Pp. 59. Thirty-five cents. Number 3, Pp. 45. Twenty-five cents.

These leaflets by the professor of religious education in Boston University were issued as a study course for the guidance of the Malden City Council of Religious Education. They represent an attempt to put religious instruction upon the same systematic basis as secular instruction in the public schools. Leaflet Number 2 is especially interesting in that it brings together in brief compass accounts of the various undertakings to develop systematic religious education in all parts of the country.

HENRY EASTMAN BENNETT. *School Efficiency, A Manual of Modern School Management.* Boston: Ginn and Company, 1917. Pp. x, 374. $1.25.

There will always be a demand for a general, introductory treatise on school management. In the present book the author touches lightly on such topics as school grounds, buildings, lighting, heating, ventilation, seats and desks, apparatus, health inspection, school organization, course of study, promotions, marking systems, reports, home study, routine, lesson plans, motivation of work, correcting papers, punishments, constructive and corrective school government, the school and the community, and the rights, duties, and aims of the teacher. While the discussion of each of these themes is necessarily brief, the author is to be commended for bringing his treatment up to date and for referring to books that have been issued within the last few years. The book will be of value to those who desire a casual view of the entire field of school activities, and who do not care to go very deeply into any particular subject.

RANDOLPH BOURNE. *Education and Living.* New York: The Century Company, 1917. Pp. x, 236. $1.25.

This is a collection of papers reprinted from *The New Republic* in which the author attempts to give glimpses of new tendencies in American schools. The author says "The public school is the most interesting and the most fruitful of our American enterprises . . . . The school fascinates me because there is almost no sociological, administrative, or psychological truth that cannot be drawn out of its manifold life. It is the laboratory for human nature, and the only one that is simple enough to study with any prospect of quick enlightenment. Experiment in education has come to stay, and this means that we have it in our hands to approach ever more closely our ideal of education as living." By experiment, however, we must not think that the author has anything so definite in mind as mental or educational tests. He means rather such undertakings as the Gary Plan, the Fairhope Summer School, the Ettinger Plan, continuation schools, and the various school surveys. As an easy popular expression of new tendencies in education, the book will serve an excellent purpose.

GAMALIEL BRADFORD. *A Naturalist of Souls.* New York: Dodd, Mead and Company, 1917. Pp. 292. $2.50.

This is an interesting venture in psychography. Not the psychography of William Stern and the scientific psychologists, but psychography in the sense of Saintsbury, Sainte-Beuve, and other interpreters of literature. In the first chapter the author explains what he understands by psychography. A portrait presents only a single phase of its subject, while the psychograph seeks "out of the perpetual flux of actions and circumstances that constitute a man's

whole life to extract what is essential, what is permanent and so vitally characteristic." It differs from biography in that biography is bound to present an elaborate sequence of dates, events and circumstances to make the narrative complete, while the psychograph selects only that which is vital to the individual subject. Psychography differs from psychology in that "the latter does not deal primarily with individuals but with general principles." Among the authors visited in these diverting excursions are the poet Donne, the Italian pessimist poet Leopardi, Anthony Trollope, Richard Burton, author of *The Anatomy of Melancholy*, Alexander Dumas, the Earl of Clarendon, the younger Pliny, Francis of Sales, and the novelists of the late Greek period. The essays are delightful reading, simple, direct and unaffected, investing these well-nigh forgotten authors with a charming reality.

F. S. Breed and E. F. Down. *Measuring and Standardizing Handwriting in a School System.* Reprinted from the Elementary School Journal, Volume 17, Number 7. March, 1917. 470-484.

The authors contend that each school system should measure its present achievement in handwriting, should establish norms of achievement, and should construct its own scales for measuring achievement. The present article deals with the construction of such a scale on the basis of specimens of handwriting from grades III–VI of the Highland Park, Michigan, schools.

Ernest R. Breslich. *Third Year Mathematics for Secondary Schools.* Chicago: The University of Chicago Press, 1917. Pp. xviii, 369. $1.00.

Ernest R. Breslich. *Logarithmic and Trigonometric Tables and Mathematical Formulas.* Chicago: The University of Chicago Press, 1917. Pp. xviii, 118. Seventy-five cents.

This book with the accompanying tables furnishes the third unit in the author's series of texts in secondary mathematics. It continues the plan, so ably set forth in the previous volumes, of carrying all phases of secondary mathematics side by side, instead of devoting a certain time to algebra, another period to plane geometry, a third to solid geometry, and a fourth to trigonometry. To show the scope of the treatment, the first chapter presents the elements of functions of one unknown with graphic representations, chapter two deals with trigonometric functions, chapter three with linear equations in their development and philosophical significance. Subsequent chapters treat quadratic equations in one and two unknowns, factoring, exponents and radicals, logarithms and the use of the slide rule, the solution of triangles, relations between functions of several angles, the binomial theorem, areas of surfaces, computation of volumes, and polyhedral angles. The plan commends itself strongly as well-balanced and carefully worked out. From a general point of view it has two marked advantages. It makes it possible

to give all pupils a fairly broad view of the whole subject of mathematics, and then allow them to drop the study if they are not mathematically inclined. For the pupils who go on in the subject it organizes and binds together the various parts in such a way as to make more likely their subsequent functioning in any practical activity demanding mathematical skill.

SARA CONE BRYANT. *Stories to Tell the Littlest Ones.* Boston: Houghton, Mifflin Company, 1916. Pp. xiv, 178.

Nothing appeals to children like a story and in this volume are to be found many happy efforts of a veteran story-teller. The book will be a boon to mothers and teachers.

C. MACFIE CAMPBELL. *The Sub-normal Child—A Study of the Children in a Baltimore School District.* Reprinted from Mental Hygiene, Volume 1, Number 1, January, 1917. 96–147.

An intensive study of the Locust Point District in Baltimore, comprising over 1500 children of school age. The sub-normal children have been studied in detail and tables are given showing chronological age, mental age (Binet-Simon), pedagogical age (grade), heredity and other data which throw light on the case.

FRANK G. CARPENTER. *Around the World with the Children.* Cincinnati: American Book Company, 1917. Pp. ix, 133.

This is an entertaining and magnificently illustrated introductory book to geography. It will acquaint the children with the peoples of distant lands, their customs, mode of life, and occupations, particularly children's occupations.

WILLIAM ESTABROOK CHANCELLOR. *A Theory of Motives, Ideals and Values in Education.* Boston: Houghton Mifflin Company, 1917. Pp. xiii, 534.

This is a reprint of a book first published ten years ago, and rather well known in the educational world. There is no evidence of any revision. The author discusses in an entertaining way all sorts of educational topics, and manifests withal such a wealth of literary allusion and citation, such a depth of sociological and philosophical erudition that he cannot fail to make an impression upon the reader. There is much good advice (perhaps the carping critic might charge the author with inordinate preaching), and there are many passages that stimulate to reflection, perhaps to opposition. Certainly a perusal of the book will result in no little gain to the reader.

BERTHA M. CLARK. *Laboratory Manual for Introduction to Science.* Cincinnati: The American Book Company, 1917. Pp. 203.

This loose-leafed laboratory note book is to accompany the author's text on Introduction to General Science and gives detailed directions for a large number of simple and practical experiments.

George Albert Coe. *A Social Theory of Religious Education.* New York: Charles Scribner's Sons, 1917. Pp. xiii, 361. $1.50.

"The future of society depends upon the sort of social education that we think it worth while to provide. All the plans and methods of religious education have now to be recognized with reference to social relations and experiences. A new measure is provided for the material that goes into the curriculum of instruction. The educational relations between church and state have a different look when we approach them from the standpoint of a thoroughly socialized religion. Not less true is it that emphasis now shifts from one part of educational psychology to another." A careful examination of the Bible from the politico-social point of view shows that its entire scheme is autocratic. But at the present time we are fighting autocracy with all our might. Can a democratic religion be worked out on a Biblical basis? The author believes that it can, and tries to show how it is to be done. The book should be carefully considered, especially by those who have their doubts about the social outcome of any form of religious education.

H. E. Conard and G. F. Arps. *An Experimental Study of Economical Learning.* Reprinted from the American Journal of Psychology, Volume 27, October, 1916. 507–529.

The problem of this experiment was to determine the loss in speed and accuracy accompanying the traditional methods of teaching the four fundamental operations in arithmetic as compared with more economical methods. On the basis of scores in Courtis Arithmetic Tests, series B, 64 first year high school pupils were divided into two equal groups, one of which was trained by traditional methods, the other for an equal length of time by the improved method. The Courtis Tests were then given a second time in order to determine the exact amount of superiority attained by the economical method. Elaborate tables and graphs of the results are shown, and a superiority of approximately 33 per cent. is claimed for the economic method.

Katherine M. Cook and A. C. Monahan. *Rural School Supervision.* Bulletin 1916, Number 48. Washington, Bureau of Education, 1917. Pp. 63. Ten cents.

The monograph presents a sketch of the origin and development of school supervision, both state and local, recounts the present status of rural supervision, and describes special forms of rural supervision in various parts of the country.

Charles William Dabney, Chairman. *A Study of Educational Conditions in Mexico, and an Appeal for an Independent College.* Cincinnati: Committee for the Study of Educational Conditions in Mexico, 1916. Pp. 93.

This is a report on educational conditions in Mexico by a committee of prominent educators in this country. The author of most of the report is Professor George B. Winton, now of Vanderbilt University, for thirty years a teacher in Mexico. The report is an excellent history of Mexican education, a survey of the present situation, and proposals for betterment.

CHARLES A. ELLWOOD. *An Introduction to Social Psychology.* New York: D. Appleton and Company, 1917. Pp. xiii, 343. $2.00.

In this book the author endeavors to bring to bear upon the problems of society the methods and results of modern psychology. He believes that the explanation of social phenomena is to be sought in the underlying traits and dispositions of the individual, in the influences of the environment which act upon his plastic nature, and in the resultant aims and standards which he develops. The first three chapters lay the foundation for the discussion in the development of the biological point of view, and in an outline presentation of the salient features of human and comparative psychology. The body of the work deals with the nature of social unity, social continuity, and social change under normal and abnormal conditions. Then follow applications of social continuity and change in such phenomena as instinct, imitation, sympathy, and social progress. The book is sane, reasonable, and carefully thought out, and is much freer from the vagaries of faculty psychology (such as the atomistic conception of instinct as an all-explanatory agency) than most recent works on social behavior.

HERMAN FERNAU. *The Coming Democracy.* New York: E. P. Dutton and Company, 1917. Pp. viii, 321. $2.00.

There are some Germans whose thought has advanced beyond the stage of medieval feudalism,—but they dare not live in Germany. The author of this book lives in Switzerland, and those who have read or heard of his earlier work, "*Because I am a German,*" know that it would not be safe for him to show himself in Germany at present. This book is a strong appeal to the German people to rid themselves of their fetishistic adoration of autocracy, and to assert their right to govern themselves. The first chapter gives a very careful survey of the causes of the war, showing conclusively that hostilities were deliberately brought on by the German war party. There follow chapters on the basis of dynastic power, with special reference to the Hohenzollern dynasty, the principles of German policy, the subordination of German "Kultur" in the service of dynastic interests, the conception of the "Vaterland" as developed by the imperialists, the underlying forces which animate the nations involved in the present conflict, and the book closes with a stirring call to the German people to rouse themselves from their hypnotic trance and press onward to democracy. It is a splendid book, and is to be cordially recommended to those who desire to understand the psychology of the German people at the present time.

Eleanor A. McC. Gamble, Editor. *Wellesley College Studies in Psychology, No. 2.* Psychological Monographs, Vol. 22, No. 4, 1916. Whole No. 96. Pp. 192. $2.00.

This volume contains four studies, three of them dealing directly with the learning process. The first study is by Dr. Helen D. Cook and Miss Florence M. Kunkel, and brings out the fact that the contrast colors differ from the complementary in the case of the primary colors red, yellow and blue. For green the contrast and the complementary are approximately the same. The second study, conducted by Professor Gamble and Miss Wilson, showed that both in learning and in recall of nonsense syllables associations of spatial position play an important role. From the third study, by Professor Gamble, and the fourth, by Miss Josephine N. Curtis, it would appear that each learner has an optimum rate of learning, which differs with different individuals, and which gives the best results both in tenacity of retention and in exemption from fatigue.

Kate Gordon. *Educational Psychology.* New York: Henry Holt and Company, 1917. Pp. vi, 294.

In this book we have another attempt "to apply the methods and facts known to psychology to the questions which arise in pedagogy." The author tells us that an elementary knowledge of psychology is presupposed, though why this should be necessary apart from the regulations of certain institutions it is difficult to see. Certainly the text is not so involved as to require such a preliminary study for its mastery. A glance over the chapter headings leads one to suspect that the author, as is customary in such attempts at "application," has not got very far from the traditional course in introductory psychology. We find such rubrics as growth, instinct, motor capacities, sensory capacities, learning, observation, memory, reasoning, the syllogism, transfer of training, attention, feeling, and will. The last three chapters deal with language, drawing and arithmetic, but no mention is made of the recent experimental work in the measurement of attainments in these subjects. Many interesting class exercises and well-chosen illustrative materials make the book valuable to the teacher of psychology, but it is rather irritating to find a writer on "educational psychology" so completely ignoring the rapidly growing science of educational measurements.

Maximilian P. E. Groszmann. *The Exceptional Child.* New York: Charles Scribner's Sons, 1917. Pp. xxxiii, 764. $2.50.

The author uses the term "exceptional" child in its broadest sense to include all types of deviation from the "average," whether up or down. He tries to give a perspective of the entire situation, including questions of heredity and family history, of environment and social-economic conditions, of child hygiene and public sanitation, of medical inspection and clinical work, and of psychologic and psychopathic investigation. There is a chapter on the normal

child and another on exceptionally bright children, and the remainder of the discussion is taken up chiefly with the problems presented by those below the normal. In the chapter on Binet Tests no mention is made of the Stanford revision, and the entire attitude is one of negative criticism. In the list of tests used by the author we find the customary ear and eye tests, color perception, visual and auditory memory, tactile, smell and taste tests, motor coordination, sense of location, balance, controlled association, imitation (Knox Cubes), concentration, language, puzzles, ciphers, reading, writing, Courtis arithmetic tests, Healy form boards, drawing and construction tests. It is obvious that these are applicable only to individual, not to group testing. An appendix of 200 pages contains a medical symposium on a typical children, and a 27 page bibliography.

H. L. Hollingworth and A. T. Poffenberger, Jr. *Applied Psychology*. New York: D. Appleton and Company, 1917. Pp. xiii, 337. $2.25.

For centuries psychology was scarcely more than metaphysical speculation about the way we come to know things. In the last generation it appropriated experimental methods and became a descriptive, if somewhat formal, science. Now the advance outposts proclaim that the function of psychology is to enable us to foretell what response an individual with a given heredity and a given previous training will make to a given situation. From this point of view all psychology is applied psychology. In the present volume, however, the authors are scarcely so ambitious, but select for their special consideration such topics as heredity, efficiency and learning, environmental conditions of efficiency, work, rest and fatigue, drugs and stimulants, business and factory management, the psychology of the worker, and psychology in its relation to salesmanship, law, medicine, social relief, and education. It is more of a program than a handbook of scientifically ascertained facts, but it brings together in convenient form and in the scientific spirit the scattered and fragmentary data that have been discovered in these fields, and thus will contribute materially to the further development of the science.

F. Wood Jones. *Arboreal Man*. New York: Longmans, Green and Company, 1916. Pp. x, 230. $2.40.

This is an extremely interesting account of an important stage in man's history. It is well established that man came from an arboreal stock. Accepting this view the author raises the further question when did this stock become arboreal, and when did it give rise to a creature which we could possibly call human. In masterly fashion he marshals the evidence from comparative anatomy and physiology bearing upon the question, and gives a vivid and detailed picture of the way in which this change of habitat was carried out, and the resulting consequences upon the evolution of animal structure. Here we see how the fore-limbs became emancipated from the burden of sustaining weight, and gradually developed that mar-

vellous flexibility and deftness to which civilization owes so much in the work of the human hand. The recession of the snout and jaws, the reduction of the teeth, the development of the face and cranium, the poise of the head, the changes in the viscera and in respiration due to the erect position, the enlargement of the brain, the shifting of sensory controls from touch and smell to sight and hearing, and the higher developments of cerebral functions are described step by step, with a welath of pictorial illustration which makes the evidence overwhelming. Such a book is profoundly educative, in that it gives us a better understanding of our origin, and deserves to have the widest possible circulation amongst the educational public.

J. ALBERT KIRBY. *The Kirby Rhythmic Method of Penmanship.* New York: Newson and Company, 1916. Pp. 133.

This writing method is designed to meet the needs of the modern business world. The author is a teacher of penmanship in the High School of Commerce, New York City, and in working out his method has given particular attention to the two requisites of modern business and professional writing, legibility and ease of execution. These are to be attained through emphasis on rhythmic movement in accordance with the developing action and habit systems of the child. The teacher's manual and the accompanying sets of practice pads take account of three natural divisions in the development of writing habits, primary, intermediate, and advanced, depending upon the age and muscular control of the child. The method would seem to commend itself as a logical and well-planned means for securing easy, comfortable, speedy, and legible writing. It is interesting to note that phonograph records have been made to assit teachers in the development of the rythmic idea.

WILLIAM MCANDREW. *The Public and the School.* Yonkers-on-Hudson: World Book Company, 1916. Pp. xii, 76.

This is a new type of school report, and one that deserves to be widely known and imitated. In the first place the report is made in one hundred and eleven pithy and pungent paragraphs. The author has boiled down what he wants to say, and says it directly, effectively, even brilliantly. In the next place the report is illustrated,—not with the graphs and diagrams which we are beginning to associate with statistical reports, but with hand drawn illustrations, hitting off salient features of the discussion in a manner that would do credit to a cartoonist on one of our great newspapers. Finally on almost every page the author avows his faith in testing and measuring results. Not only is it a matter of faith but likewise of works. Accounts are given of ingenious quantitative tests in arithmetic, penmanship, spelling, composition, grammar, and following directions. The finding of Courtis that New York school children rank low in arithmetical accuracy but high in speed is caustically

interpreted, "It takes us less time to get a thing wrong here than it does in the average school system," and is pictorially represented by a huge loaded dray dragged at top speed through the street by straining men and women but with the bundles dropping off in a stream behind. Throughout the report the author advocates efficiency standards, efficiency records, and the recognition of efficient attainments by corresponding rewards. Many of his tests deserve further application and standardization.

FREDERICK WILLIAM HUGH MIGEOD. *Earliest Man.* New York: E. P. Dutton and Company, 1916. Pp. xii, 132.

This admirable little book sketches the evolution of man from his simian ancestry down to the time when he attained the rank of *Homo primigenius*. "The argument is based on his mentality at different stages. Inquiry is made whether his brain is sufficiently developed at a given epoch to admit of its being assumed that he could act in a certain given way, and, inversely, what mental development he must have had when he could manufacture a certain chipped stone." The various chapter headings indicate the scope of the work: The Dawn, Primary Instincts, Proto-man, Progress in the Arts, From Eoliths to Palaeoliths, Origin of Speech, Social Organization, and Appendices giving the cranial capacity of different animals, a chronological table of development, and a hypothetical descent of man. The book gives a splendid picture of the probable steps in the evolution of human civilization, and makes one realize the tremendous reaches of time involved compared with which our historic time is a mere fragment.

WILLIAM HENRY PYLE. *The Science of Human Nature: A Psychology for Beginners.* Boston: Silver, Burdett and Company, 1917. Pp. vii, 229.

There is some warrant for the assumption that the number of text-books and treatises issued in a given department of human knowledge is a fair index of the progress of intellectual activity in that field. According to this criterion there is no doubt that psychology is in a flourishing condition. The present text is designed especially for students in high and normal schools. There is no doubt that a study of human nature would be much more valuable for high school students than many of the subjects on which they spend (or waste) their time. Carefully planned courses in literature, history, government, psychology, and sociology would afford a far more adequate foundation for actual living than the grind on languages and mathematics with which the course of study has been loaded. In this book the author has made such a selection from the entire field of psychology as he thinks will be of value to young people. The general rubrics are the familiar ones of mind and body, inherited tendencies, feeling, attention, habit, memory, thinking, etc., but the treatment is in easy narrative form, the illustrations

are drawn from familiar, every-day situations, and each chapter is followed by a summary, a list of simple class exercises, and a few carefully selected references for further reading. The last two chapters deal with individual differences and applied psychology, and put the student in touch with interesting current developments in experimental psychology.

HAROLD O. RUGG. *Statistical Methods Applied to Education.* Boston: Houghton Mifflin Company, 1917. Pp. xviii, 410. $2.00.

In his introduction the editor well says, "Such a volume forms an interesting exhibit of the progress at present being made in the organization of instruction in the subject of education. Two decades ago there would have been almost no use for such a volume, as we had not then begun to make any accurate measurements of the products of our educational efforts. Only the most general terms were then in use, while today the demand is for quantitative expression in commonly used terms which students can understand." The author's aim is to make clear to school administrators whose training in mathematics has not gone much beyond the high school the statistical and mathematical treatment of school records and experimental studies of school attainments. The introductory chapter on the use of statistical methods emphasizes the importance of the graph, and gives many illustrations of the plotting and interpretation of graphs. The various sources of educational data are next considered, and methods and mechanical devices for tabulating these data are extensively described. The statistical classification of data with reference to attributes and variables is then discussed, and the terms "scale," "unit," "class-interval," and "frequency distribution" are explained and exemplified. The significance of such indications of central tendency as the "mode," the "median" and the "mean" is made clear, measures of variability as range, quartile deviation, standard deviation and mean deviation are expounded, the usefullness of the frequency curve in educational measurements is emphasized, and there is an extended treatment of the various indices of correlation. The book will not supplant Thorndike's pioneer work "Mental and Social Measurements," but on account of the detailed explanations and the richness of illustration will be a valuable supplement to it.

WALTER ROBINSON SMITH. *An Introduction to Educational Sociology.* Boston: Houghton Mifflin Company, 1917. Pp. xvii, 412. $1.75.

The study of social institutions is growing steadily in favor, and the recognition of education as one of the most important of these institutions is an indication of the drift of public sentiment. We have not yet reached clearness in the fundamental aims of education. The old competitive individualism is gradually giving way to the conception of the individual as a member of a social group,

and ideals of group welfare are becoming more and more dominant in our educational policies. It is well to have this systematically set forth in a volume on educational sociology. The first ten chapters of the book sum up the essential principles of sociology that are significant for education. There is a discussion of social organization, of the individual and the social group, of such primary social groups as the family, the play group, and the community, of intermediate groups as the labor unions, clubs, scientific and literary societies, and religious organizations, of the relation to the state and to democracy, and of the evolution of the modern school. In part II we have the educational applications of these fundamental principles. Social and educational surveys, social aspects of school administration, the socialization of discipline and the curriculum, the social basis of utilitarian and vocational education, cultural aspects of a socialized education, and the socialization of teaching methods are the chief topics considered. The book will have a stimulating influence on educational thought.

LEWIS M. TERMAN, and Others. *The Stanford Revision and Extension of the Binet-Simon Scale for Measuring Intelligence.* Educational Psychology Monographs, No. 18. Baltimore: Warwick and York, 1917. Pp. 179. $1.40.

This monograph presents the detailed, long-awaited account of the Stanford Revision. There is a brief history of the studies on which the revision is based, a list of the tests used, an examination of the distribution of intelligence for different ages, a consideration of the rate of growth and the validity of the intelligence quotient, a study of sex differences, an estimation of the relation of intelligence to social status and to school success, a critique of the validity of the individual tests, and a discussion of the fundamental principles involved in the formation of a scale of intelligence. The scale yields a median intelligence quotient of approximately 100 with unselected children of any age from five to fifteen. The expression of mental age in absolute years has no significance and should be abandoned in favor of the intelligence quotient. This remains fairly constant for a child of any grade of intelligence until well into adolescence, and there is no evidence of "nascent stages" or "adolescent spurt." The girls seem to be slightly superior to the boys at each age up to thirteen, although in some tests the boys are uniformly superior. There is nothing to support the widespread belief that extremes of intelligence are more common among boys than among girls. There is a marked correspondence between social status and intelligence, and this seems to rest upon original endowment rather than upon environment. The correlation between teachers' rankings and intelligence quotient was .48, an index high enough to warrant more extensive use of the scale in school diagnosis and classification.

VOLUME IX MARCH, 1918 NUMBER 3

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

A REVISED DIRECTIONS TEST

RUDOLPH PINTNER AND HERBERT A. TOOPS

Ohio State University

The directions tests of Woodworth and Wells* have proved themselves so valuable and are being used so extensively by various psychologists that it is surprising that no norms have yet appeared. The writers offer the data they have collected up to the present time in the hope that they may prove serviceable to other workers.

The Easy Directions Test B and the Hard Directions Test have been used. These tests are reproduced in Figures 1 and 2 respectively. The numbering of the directions does not appear on the test sheets. We have inserted the numbers on the figures for purposes of reference. The Easy Directions Test was given to 429 children ranging in age from 6 to 16. The Hard Directions Test was given to 1327 children ranging in age from 7 to 17, and to a heterogenous group of adults.

The instructions in each case were very simple: "I want to see if you can fill these out according to directions. Do exactly what each line tells you to do. Try to fill them out correctly." No definite time limit was adhered to because we wished the children to try all the directions. As a matter of fact the majority of pupils finished in much less than ten minutes. The test was given as a class test.

A score of *one* was given for each direction correctly answered in the Easy Directions. The total possible score in this test is therefore 20. A direction had to be literally correct before a score was allowed. Direction 1 was given credit only when some sort of a mark crossed out the g and the g only. Direction 2 received credit only for the figure "2" placed between the two dots, the word "two" not receiving credit. Direction 3 received credit for the answer, "three," "3", or "III," "three feet," etc., or any answer conveying

*WOODWORTH, R. S. AND WELLS, F. L. *Association Tests,* Psychol. Monog., Vol. XII, No. 57, 1911.

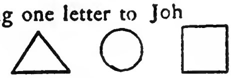
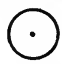
1. Cross out the *g* in *tiger*
2. Write 2 between the two dots: • • ——
3. How many feet make a yard?
4. Write + over the longest word: It rained yesterday.
5. Put a dot below this line: ————————
6. Write the sum of these numbers: $\frac{3}{4}$
7. Make a boy's name by adding one letter to Joh
8. Make a cross in the circle: 
9. What comes next after D in the alphabet?
10. Write 7 in the largest square: 
11. Cross out the blackest letter in TE**X**AS
12. Write *g* on the egg-shaped figure: 
13. Make two dots between these lines: 
14. Put the sign = where it belongs: 3 + 2 5.
15. Write here.....the middle letter of *get*.
16. Put a nose on this face: 
17. Add a cross and make these rows equal: X X X / X X X X
18. Put a dot in the circle, below the center: 
19. Draw a line around the three dots: • • • • • • •
20. Cross out the last word in this sentence.

Figure 1. Easy Directions. Test B.

1. With your pencil make a dot over any one of these letters F G H I J, and a comma after the longest of these three words: boy mother girl. 2. Then, if Christmas come, in March, make a cross right here.........but if not, pass along to the next puestion, and tell where the sun rises............ 3. If you believe that Edison qiscovered America, cross out what you just wrote, but if it was some one else, dut in a number to complete this sentence: "A horse has......feet." 4. Write *yes*, no matter whether China is in Africa or not......; and then give a wrong answer to this question: "How many days are there in the week?"............. 5. Write any letter except *g* just after this comma,......and then write *no* if 2 times 5 are 10...... 6. Now, if Tuesday comes after Monday, make two crosses here......, but if not, make a circle here......or else a square here...... 7. Be sure to make three crosses between these two names of boys: George......Henry. 8. Notice these two numbers: 3, 5. If iron is heavier than water, write the larger number here........, but if iron is lighter write the smaller number here........ 9. Show by a cross when the nights are longer: in summer?......in winter?...... 10. Give the correct answer to this question: "Does water run uphill?".......... and repeat your answer here........ 11. Do nothing here (5+7=......), unless your skipped the preceding question; but write the first letter of your first name and the last letter of your last name at the ends of this line:

FIGURE 2

the idea, three. No. 4 received credit only for the sign + written *above* "yesterday," no credit for "X" or for any other sign or in any other position than *above* "yesterday." No. 5 had to have the dot fairly below the black line. Credited responses to directions 6 are "7" or "seven" *with* or *without* drawing the addition bar, and the answer placed in any position with respect to the two figures to be added. In direction 7, credit was given when "n" or "N" was added to the end of "Joh," but no credit was given for writing out the word "John" and neglecting to complete the partial "Joh." No. 8, credit for either "X" or "+" in the circle. No. 9, credit for "e" or "E". No. 10, credit only for the figure, "7," not for the word, "seven," in the largest square. No. 11, credit for any mark through "X" only. No. 12, credit for "G" or "g" (perhaps a slight deviation from our principle of literal exactness). No. 13, credit for two dots placed anywhere between the two lines. No. 14, credit only for "–" placed in the proper place, no credit for "–" or "+," etc. No. 15, credit for "e" or "E." No. 16, credit for any mark filling up

the greater portion of the blank left in the outline. No. 17, credit for "X" in the proper place. No. 18, credit for a dot anywhere in the lower one-third area of the circle below the dot in the centre. No. 19, credit for any shaped line enclosing the three dots only. No. 20, credit for a horizontal line through "sentence" or for any mark crossing out *at least two* of the letters, no credit for any line crossing out only one letter, as in "sentence," because this is a response to a misinterpretation of the direction whereby "letter" is understood instead of "word."

The Hard Directions Test was scored in the manner suggested by Woolley.* A credit of one point was given to each simple direction and a credit of two points for each compound or conditional direction answered correctly. If the conditional direction was only partially correct then a credit of one point was allowed. The conditional directions are numbers 2, 3, 6, 8, and 11. *No credit was given either part of the condition unless one or the other of the parts was attempted as shown by some mark or other on the paper.* This procedure was necessary because credit is sometimes given for doing nothing in one part of the condition. First direction, one credit for a dot *above* any *one* of the five letters and one credit for a comma after "mother," thus making two credits if both are correctly done. Second, one credit for *not* making a cross, and one for "east," no credit for "west," etc.; if a cross is made in the first part of the direction and nothing done to the second part, then one credit was given (an intelligent response based on wrong information); if both parts are answered, one credit was given for answer to the second half. Third, credit for *not* crossing out anything in the *preceding* sentence and then if nothing is crossed out an additional credit for "four" or "4" placed in proper space; one credit is given if the second half is answered correctly and the first part incorrectly. This direction is unsatisfactory, since the subject may have already crossed out the word written in the preceding sentence before coming to this direction because the word was wrong. Fourth, one credit for "yes" and one credit for *any* number or figure other than "7" or "seven." Fifth, one credit for *any* letter other than "g," and one credit for "no." Sixth, credit for "XX" or "+ +" and an additional one for *not* doing anything in the other two spaces; a single credit of one is given for filling in the second part of this direction

*The writers wish to thank Mrs. H. T. Woolley for sending them detailed directions as to her method of scoring.

without filling in the first (an intelligent response based on wrong information). Seventh, one credit for "XXX" or "+++" or "***." Eighth, credit of one for "5" in first blank and an additional credit if nothing is done in the second, one credit if the second part *alone* is filled with "3;" a credit of one if "5" is placed in the first blank and "3" also in the second. Ninth, one credit for "+" or "X" after "winter," no credit for a mark after "summer" or for a mark after both "summer" and "winter." Tenth, one credit for "no" or "it does not," etc., and one additional credit for repeating *exactly* the first answer; in case the first answer were "yes" or anything else, and this is correctly repeated, one credit is given for the correct repetition. Eleventh, one credit for doing nothing to the first part *if* the preceding direction was attempted (one credit for "12," if the preceding direction was not attempted) and one credit for "J. e." or "J. E.," if the subject's name is John Doe, placed after the word, "line." In accordance with the principle first laid down, no credit at all is given for this direction (if the preceding direction was attempted) unless the second part of the direction is attempted.

The results for the individuals tested grouped according to age are given in Tables I and II. Table I shows the percentiles at each age for the Easy Directions Test. The table is to be read as follows:

TABLE I

Easy Directions Test. Percentiles

Age	6	7	8	9	10	11	12	13	14	15	16
Percentiles											
100	8	16	16	20	20	20	20	20	20	20	20
90	8	13	12	19	20	19	20	20	20	..	..
80	6	10	12	19	20	19	20	19	20	..	..
70	2	8	10	19	19	19	19	19	20	..	..
60	2	6	9	18	19	19	19	19	19	..	..
50	2	4	7	18	18	18	18	19	19	19	20
40	1	3	7	18	18	18	18	18	18	..	..
30	0	2	5	17	18	17	17	18	18	..	..
20	0	1	3	17	17	17	17	17	18	..	..
10	0	0	1	14	17	15	16	17	17	..	..
0	0	0	0	14	15	5	12	15	14	9	13
No. of Cases	11	81	32	10	38	60	69	58	38	24	8

the highest or 100 percentile score made by six-year-old children is 8; ten per cent. of the children below the highest scored 8 also, and so on down the column. The other columns are to be interpreted in the same way.

Table II shows the percentiles at each age for the Hard Directions Test. The table is to be interpreted in the same manner as Table I.

TABLE II

Hard Directions Test. Percentiles

Age / Percentile	7	8	9	10	11	12	13	14	15	16	17	Ad.
100	18	19	20	20	20	20	20	20	20	20	20	20
90	14	16	18	18	18	19	19	20	20	19	19	20
80	13	13	17	17	17	18	19	19	19	19	19	20
70	11	12	16	16	16	17	18	18	18	18	16	20
60	8	10	13	13	15	16	18	18	18	17	16	20
50	8	9	12	12	14	14	17	17	16	16	16	19
40	6	8	11	11	13	13	16	16	15	16	15	19
30	5	7	9	8	10	11	15	15	14	15	13	18
20	3	6	8	7	9	10	13	14	12	13	10	17
10	1	3	6	6	3	6	9	11	9	9	8	15
0	0	0	0	0	0	0	0	0	0	0	8	11
No. of Cases	22	79	115	124	124	154	213	253	144	45	10	44

The table for the Easy Directions Test shows that the test discriminates fairly well between the varying abilities of six, seven and eight-year-olds, in-as-much as all these ages show a gradual increase in the scores from the 0 to the 100 percentile at each age, and, furthermore, each age shows an increase over the preceding age. The number of cases tested at age nine is not sufficient, but from that age on we note very little difference in the percentile scores for each age and scarcely any increase from age to age. The Hard Directions Test shows discriminative capacity from age seven up to age thirteen, or fourteen, in-as-much as the percentile scores show a gradual increase for each age and also from one age to the other.

The fifty percentile at each age gives the median or norm for each age. Figure 3 shows the median scores for both the Easy and the Hard Directions Tests. We note that on Test B the scores increase gradually from age six to age eight, and that at age nine there is a sudden increase in score with but little increase thereafter. This would seem to indicate that the directions are entirely too easy for all ages after age nine. The Hard Directions increase gradually from age seven to thirteen with no increase thereafter except for the adults. This would seem to indicate that this test is too easy after age thirteen. The peculiar overlapping of the curves at ages seven

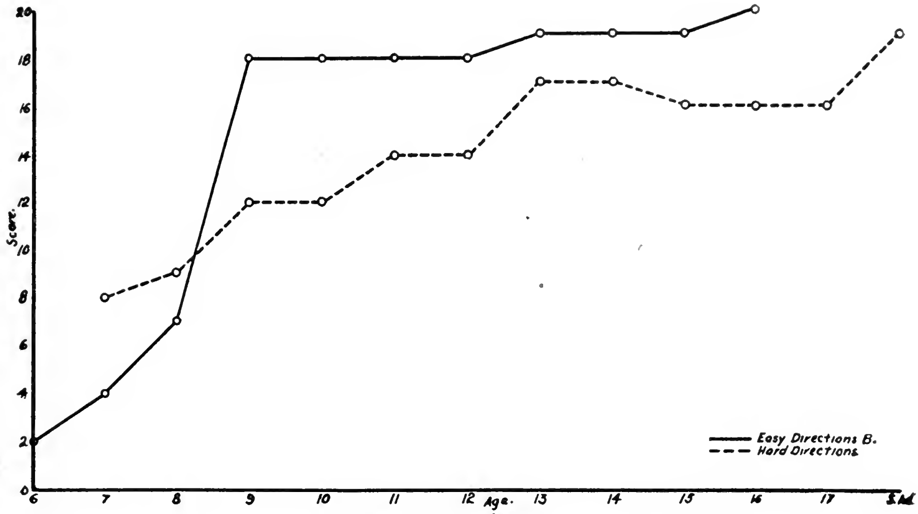


Figure 3. Age Medians.

and eight is due to the fact that the seven and eight-year-olds tested on the Hard Directions Test are mainly the brighter accelerated pupils found in the higher grades. These norms are obviously too high for seven and eight-year-olds in general.

If these tests are to be used, then, they can only be used within the limits already mentioned, Test B from age six to age nine inclusive and the Hard Directions from age seven or eight to age thirteen inclusive. It would be much more desirable to have one test covering all ages from age six to age thirteen, or even better, to adult. The two tests might simply be massed together as one test giving us a test covering ages six to thirteen. This procedure would, however, give us too cumbersome a test because much useless material, due to relative equality of many of the directions, would evidently be included. It, therefore, would seem to be a better procedure to determine the relative difficulty of the different directions and eliminate those found inadequate.

Relative Difficulty of the Directions

Easy Directions Test. From a mere inspection of the directions in either test, it is obvious that they differ greatly among themselves in difficulty. From the data gathered we are able to estimate the relative difficulty of each direction.

TABLE III

Percentage of Each Age Passing Each Direction. Test B.

Direc. No.	Chronological Age 7	8	9	10	11	12	13	14	15	16	(a) Total	(b) Aggregate %	(c) Aggregate % ranking in order of difficulty
1	22.1	21.9	70.6	92.5	90.0	88.5	93.0	92.0	95.8	100.0	72.2	766 4	13
2	34.6	40.6	76.5	85.0	93.3	90.0	88.0	92.0	87.5	100.0	75.1	787.5	9
3	6.2	9.4	58.8	75.0	80.0	87.0	88.0	94.7	87.5	100.0	63.5	686.6	17
4	3.7	0.0	41.2	65.0	66.7	75.5	63.8	65.0	70.9	71.4	49.6	523.2	20
5	50.5	62.5	88.5	97.5	93.3	94.3	93.0	100.0	95.8	100.0	84.0	875.4	3
6	3.7	3.1	41.2	80.0	85.0	78.3	91.4	86.7	54.2	71.4	59.1	595.0	19
7	16.0	25.0	58.8	92.5	96.7	91.4	82.7	94.7	95.8	85.7	70 8	739.3	15
8	11.1	18.7	88.5	95.0	86.7	91.4	94.9	100.0	95.8	100.0	71 8	782.1	12
9	22.2	31.2	70.6	92.5	96.7	97.1	96.5	94.7	100.0	85.7	76.0	787.2	10
10	60.5	78.3	94.5	97.5	96.7	97.1	98.3	100.0	100.0	100.0	89.2	922.9	2
11	42.0	62.6	82.6	97.5	91.7	97.1	100.0	100.0	100.0	100.0	83.3	873.5	4
12	83.4	53.1	82.6	97.5	93.3	97.1	94.9	97.5	91.6	100.0	80.0	841.0	6
13	29.6	34.4	70.6	90.0	90.0	94.3	89 6	92.0	100.0	100.0	75.1	790.5	8
14	39.6	50.0	70.6	87.5	90.0	92.8	98.3	94.7	87.5	100.0	78.4	811.0	7
15	12.7	18.7	70.6	82.5	95.0	92.8	94 9	97.5	91.6	100.0	71.1	756.3	14
16	64.3	78.3	100.0	97.5	100.0	98.6	100.0	100.0	95.8	100.0	90.8	934.5	1
17	34.6	40.6	82.6	97.5	98.3	98.6	96.5	97.5	100.0	100.0	80.9	846.2	5
18	17.3	28.1	70.6	97.5	96.7	97.1	98.3	100.0	95.8	85.7	75.8	787.1	11
19	27.1	50.0	64.9	85.0	76.7	74.0	82.7	84.0	70.9	100.0	66.6	715.3	16
20	29.7	34.4	47.1	72.5	70.0	72.5	84.5	89.4	83.3	71.4	63.8	654.8	18
No. Cases	81	32	17	40	60	69	58	37	24	7	Total Cases 425		

Table III shows the percentage of cases at each age able to respond correctly. The numbers in this and the following tables do not correspond exactly to those given in Tables I and II on account of exclusion of a few cases that could not be classified and to inclusion of extra cases where these had been obtained at the time of making up the different tables. The table is to be read as follows: Of the 81 seven-year-olds tested 22.1 per cent. were able to pass the first direction; 34.6 per cent. the second direction, and so on for each of the other directions and for the other ages. The next to the last vertical column, (b), shows the aggregate percentage passing each of the directions, found by adding the percentages at each age for the direction considered, and from this is derived the last vertical column, (c), which shows the ranking in order of difficulty of the twenty directions. Direction 16 has an aggregate percentage of 934.5, which is the largest value, and therefore this direction is given a rank of 1 because it is the easiest direction as determined by this method of estimating difficulty.

Table IV shows the score made on each direction at each age. The table is to be read as follows: All the seven-year-olds together

TABLE IV

Scores Made on Each Direction by Ages. Test B.

Age	Direction Number—																				No. of Cases
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
7	18	28	5	3	41	3	13	9	18	49	33	27	24	32	10	52	28	14	22	24	81
8	7	13	3	0	20	1	8	6	10	25	20	17	11	16	6	25	13	9	16	11	32
9	12	13	10	7	15	7	10	15	12	16	14	14	12	12	12	17	14	12	11	8	17
10	37	34	30	26	39	32	37	38	37	39	39	39	36	35	33	39	39	39	34	29	40
11	54	56	48	40	56	51	58	52	58	58	55	56	54	54	57	60	59	58	46	42	60
12	61	62	60	52	65	54	63	63	67	67	67	67	65	64	64	68	68	67	51	50	69
13	54	51	51	37	54	53	48	55	56	57	58	55	52	57	55	58	56	57	48	49	58
14	34	34	35	24	37	32	35	37	35	37	37	36	34	35	36	37	36	37	31	33	37
15	23	21	21	17	23	13	23	23	24	24	24	22	24	21	22	23	24	23	17	20	24
16	7	7	7	5	7	5	6	7	6	7	7	7	7	7	7	7	7	6	7	5	7
Total	307	319	270	211	357	251	301	305	323	379	354	340	319	333	302	386	344	322	283	271	425
Ranking	12	10.5	18	20	3	19	15	13	8	2	4	6	10.5	7	14	1	5	9	16	17	

scored 18 on direction 1, 28 on direction 2 and so on. This table may also be regarded as showing the total number passing each direction at each age because a score of one is allowed for each direction. The next to the last horizontal line shows the total score for each direction and from this total score is derived the last line which shows the ranking of the 20 directions in order of difficulty. Column (a), of Table III shows the percentage of the total children passing each direction as calculated from the total number passing each direction in Table IV.

Table V shows the ranking of the 20 directions in order of difficulty for each age. This table is derived from Table III, which showed the percentage passing each direction at each age.. Table V is to be read as follows: For the seven-year-olds direction 1 was the thirteenth in order of difficulty, direction 2 the sixth in order of difficulty and so on. The next to the last horizontal row gives the total rank for each direction and from this row is derived the amalgamated ranking in the last row of the table. Table VI shows the ranking of the directions by the three methods.

Correlation of the Three Different Methods of Ranking. The correlations of these three methods are as follows (Spearman's Foot Rule Method):

A with B, r = .97	A. Amalgamated Rank
A with C, r = .98	B. Aggregate Per Cent.
B with C, r = .99	C. Total Score.

TABLE V.

Ranking of Directions in Order of Difficulty, by ages. Test B.
(Based on the percent scoring on each direction.)

Age	Direction Number—																				Number of Cases
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
7	13	6.5	18	19.5	3	19.5	15	17	12	2	4	8	10	5	16	1	6.5	14	11	9	81
8	15	8.5	18	20	5	19	14	16.5	12	1.5	5	5	10.5	5	16.5	1.5	8.5	13	5	10.5	32
9	11.5	8	16.5	19.5	3.5	19.5	16.5	3.5	11.5	2	6	6	11.5	11.5	11.5	1	6	11.5	15	18	17
10	10	14.5	18	20	4	17	10	8	10	4	4	4	12	13	16	4	4	4	14.5	19	40
11	13	9	17	20	9	16	4.5	15	4.5	4.5	11	9	13	13	7	1	2	4.5	18	19	60
12	15	14	16	18	8.5	17	12.5	12.5	5	5	5	5	8.5	10.5	10.5	1.5	1.5	5	19	20	69
13	11.5	15.5	15.5	20	11.5	13	18.5	9	6.5	4	1.5	9	14	4	9	1.5	6.5	4	18.5	17	58
14	15	15	11.5	20	3.5	18	11.5	3.5	11.5	3.5	3.5	8.	15	11.5	8	3.5	8	3.5	19	17	37
15	8.5	15	15	18.5	8.5	20	8.5	8.5	3	3	3	12.5	3	15	12.5	8.5	3	8.5	18.5	17	24
16	7.5	7.5	7.5	19	7.5	19	16	7.5	16	7.5	7.5	7.5	7.5	7.5	7.5	5.7	7.5	17	7.5	19	7
Total	120.0	113.5	153.0	194.5	64.0	178.0	127.0	101.0	92.0	37.0	50.5	74.0	105.0	96.0	114.5	31.0	53.5	85.0	146.0	165.5	425
Ranking	14	12	17	20	5	19	15	10	8	2	3	6	11	9	13	1	4	7	16	18	

TABLE VI.

Ranking of Directions in Order of Difficulty by Three Methods. Easy Directions

Method	Ranking of Sentence Number																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A. Amal. Ranking	14	12	17	20	5	19	15	10	8	2	3	6	11	9	13	1	4	7	16	18
B. Aggr. Per Cent.	13	9	17	20	3	19	15	12	10	2	4	(	8	7	14	1	5	11	16	18
C. Total Score	12	10.5	18	20	3	19	15	13	8	2	4	6	10.5	7	14	1	5	9	16	17

Since C gives a higher average correlation with the two other methods than either of the other two methods, A and B, with the two remaining methods in each case, therefore C was taken as most probably representing the true order of difficulty of the directions. Correlating the ranking of each separate age, Table V, with the total ranking of method C, we obtain the results shown in Table VII.

TABLE VII

Correlation of Each Age with Method C. Easy Directions

	Age										
	7	8	9	10	11	12	13	14	15	16	Av'g.
Correlation r..........	.89	.84	.91	.85	.70	.82	.76	.72	.69	.51	.769

The high correlations for the lower ages and the gradual tendency of the correlations to decrease with increase of age is just what we should expect from a test difficult enough for only the lower ages. Pupils of higher ages may be expected to show their ability to better advantage on directions suited to their ability rather than upon directions much too easy for their age. An examination of Table III shows that the test is entirely too easy for all ages above age 9; the hardest direction for the ten-year-olds is passed by 65 per cent. of the group.

Hard Directions. The Hard Directions were subjected to the same kind of analysis as the Easy Directions. It was found that only a very small percentage of those tested made partial scores on any of the directions; and since the inclusion of these partial scores in the determination of the difficulty of these directions complicates the statistical work with little or no advantage, the scores of a perfect performance of both parts of the direction, where such occur, are the only ones included in the tables presented.

Table VIII shows the percent. passing each direction at each age, the total percentages passing each direction and the aggregate per cent. (sum of the per cents. at each age) with the aggregate ranking derived therefrom.

Table IX shows the number of persons passing both parts of each direction. The scores would, of course, be quite different, two credits in the original scoring having been given for directions 1, 2, 3, 4, 5, 6, 8, 10, 11, and one credit each for directions 7 and 9. It is obvious that the difficulty of each direction should be computed upon the number passing each direction, not the scores on each unless weighted properly, when the scores are not the same on every direction.

TABLE VIII

Per Cent Passing Each Direction by Ages. Hard Directions.

Direction	Per Cent Passing at the Age—									Total	Aggregate per cent	Aggregate ranking
	8	9	10	11	12	13	14	15	16			
1	6.9	31.1	31.1	39.6	53.8	66.0	65.0	69.4	62.0	57.3	424.9	9
2	34.5	55.5	55.5	58.4	59.0	70.5	74.5	77.8	76.3	68.1	562.0	4
3	58.6	80.0	71.1	62.5	61.1	78.5	74.8	70.8	76.3	72.3	633.7	3
4	17.3	51.1	33.4	66.6	62.3	75.0	73.1	67.8	71.5	65.8	518.1	6
5	13.8	40.0	42.2	41.6	66.5	76.3	84.2	71.6	76.3	68.7	512.5	7
6	27.7	51.1	53.4	60.5	69.5	83.5	74.8	70.1	69.1	70.1	559.7	5
7	34.5	60.0	62.2	66.6	80.1	83.0	90.3	88.5	90.5	80.9	655.7	1
8	6.9	26.6	33.4	41.6	50.6	59.2	62.0	60.8	38.1	50.2	379.2	10
9	10.4	37.8	28.8	56.3	51.7	61.5	69.2	57.7	64.3	57.0	437.7	8
10	34.5	64.3	73.4	70.9	74.8	86.5	89.3	78.5	81.0	79.9	653.2	2
11	3.5	15.5	13.4	29.2	47.5	56.8	60.3	55.5	42.8	47.9	324.5	11
No. of Cases	29	45	45	48	95	176	234	130	42	844		

TABLE IX

Number Passing Each Direction by Ages. Hard Directions

Age	Number Passing Direction—											Number of Cases
	1	2	3	4	5	6	7	8	9	10	11	
8	2	10	17	5	4	8	10	2	3	10	1	29
9	14	25	36	23	18	23	27	12	17	29	7	45
10	14	25	32	15	19	24	28	15	13	33	6	45
11	19	28	30	32	20	29	32	20	27	34	14	48
12	51	56	58	59	63	66	76	48	49	71	45	95
13	116	124	138	132	134	147	146	104	108	152	100	176
14	152	174	175	171	197	175	211	145	162	209	141	234
15	90	101	92	88	93	91	115	79	75	102	72	130
16	26	32	32	30	32	29	38	16	27	34	18	42
Total	484	575	610	555	580	592	683	441	481	674	404	844
Ranking	8	6	3	7	5	4	1	10	9	2	11	

Table X gives the ranking in order of difficulty of all directions by ages, as calculated from the percentage passing each direction at each age.

Table XI shows the ranking of the directons in order of difficulty for all ages by the three methods; A, amalgamated rank, B, aggregate per cent, C, total score.

TABLE X

Ranking of Directions in Order of Difficulty by Ages. Hard Directions

Age	Ranking of Direction Number											Number of Cases
	1	2	3	4	5	6	7	8	9	10	11	
8	9.5	3	1	6	7	5	3	9.5	8	3	11	29
9	9	4	1	5.5	7	5.5	3	10	8	2	11	45
10	9	4	2	7.5	6	5	3	7.5	10	1	11	45
11	10	6	4	2.5	8.5	5	2.5	8.5	7	1	11	48
12	8	7	6	5	4	3	1	10	9	2	11	95
13	8	7	4	6	5	2	3	10	9	1	11	176
14	9	6	4.5	7	3	4.5	1	10	8	2	11	234
15	7	3	5	8	4	6	1	9	10	2	11	130
16	9	4	4	6	4	7	1	11	8	2	10	42
Total	78.5	44	31.5	53.5	48.5	43	18.5	85.5	77	16	98	844
Ranking	9	5	3	7	6	4	2	10	8	1	11	

TABLE XI

Ranking of Each Direction by the Three Methods. Hard Directions

Method	Ranking of Direction Number										
	1	2	3	4	5	6	7	8	9	10	11
A. Amal. Rank.......	9	5	3	7	6	4	2	10	8	1	11
B. Aggr. %..........	9	4	3	6	7	5	1	10	8	2	11
C. Total Score.......	8	6	3	7	5	4	1	10	9	2	11

Correlating the three methods we find (Spearman's Foot Rule Method):

A with B, r = .97
A with C, r = .97
B with C, r = .95

On the basis of the reasoning in the case of the Easy Directions, since A gives a higher average correlation with B and C than does B or C with the remaining two, A should be the more reliable measure. The high average correlation of C, however, shows that C is a very reliable measure and we shall use this method throughout. It is obvious that any of the rankings would be reliable enough for our purpose. In this connection it is interesting to note that the inclusion of a group of 166 heterogenous adults, 11 backward pupils in the public schools, 17 professional people, 44 college students, 94 men seeking employment at a free public employment office, changed the total order of difficulty only in that, where before direction 5 preceded direction 2 in difficulty, after the inclusion of this group direction 5 followed direction 2 in difficulty.

Table XII gives the correlations of the ranking for each age (Table X) correlated with the total ranking in difficulty, method C.

TABLE XII

Correlations of Ranking at Each Age with Method C.

Age	8	9	10	11	12	13	14	15	16	Av.
Correlation	.84	.88	.88	.76	.95	.95	.97	.89	.89	.89

Harder Directions. Now since the two tests separately or in combination test only up to and including the thirteenth year, it seemed desirable to secure some harder directions to be included as a test for fourteen-year-olds and beyond Accordingly eight directions were secured, five of which were discarded as unsuitable. Of the three retained, numbers 5 and 8 were adaptations from Kelley's Reading Tests.*

The procedure was again the same as in the case of the last two tests. It is not necessary here to go into the details of the several operations as the norms secured from the test as a whole were inadequate, lacking differentiation; all but four of the tests were found too easy for the ages above 13; three of these four are incorporated in the final series and the fourth was excluded because of lack of differentiation from age to age and also because of equal median difficulty with one of the three chosen. The three ultimate-

*Kelley, F. J., *The Kansas Silent Reading Tests,* Jour. of Educ. Psych. Vol. VII, 2, Feb. 1916., Pp. ?3–80.

ly chosen are shown in Figure 4 as incorporated in our New Directions Test. (Their original numbers in the harder Directions Test were 8, 5, and 6, respectively).

The percentages passing each of the three chosen directions at each age are shown under "Harder Directions" in Table XIII. Direction 6, which becomes the last of our new series, *i. e.*, No. 15, proved too difficult for all but college students. Hence it was included as a test for "superior adults."

A Suggested New Directions Test

By simply combining Test B and the Hard Directions on one sheet and securing norms for the combined sheet we might thereby get a test for ages 6 to 13 inclusive; similarly, by combining this new test with the Harder Directions Test we might get a test, more or less adequate, for all ages, 6 to "superior adult" inclusive. Such a test would, however, be cumbersome, due to the inclusion of too many directions of relatively equal difficulty. We propose therefore to eliminate the useless and inadequate directions and combine those remaining into one test which shall not be cumbersome and shall differentiate between all ages from 6 to "superior adults."

The directions selected should vary in difficulty from the easiest to the hardest. The directions should so vary in difficulty that each age will show an increase in score above the preceding age. The directions selected should show increasing percentages passing each direction with increasing age. To this end they should not involve more than the simplest kind of knowledge for their carrying out. Each age should be represented in the test with one or more directions passed by the median child, or by a safe margin of that age above 50 per cent., and which is not passed by the median child of the age below, by a safe margin below that age.

With all these points in mind we have chosen 15 directions from the three tests as best fulfilling these conditions. Table XIII gives the selected directions with the percentages passing at each age as taken from Tables III and VIII and the results from the Harder Directions Test. The percentages on the directions from the several tests are based necessarily on different numbers depending on which test they originally came from. The directions for the ages, 6, 7, 8, 9, are based on 425 persons tested by test B; the directions for ages 10, 11, 12, 13, are based on 844 persons tested by the Hard Directions Test; and the directions for ages 14, 15, and Adult are based on 698 persons tested on the Harder Directions Test.

TABLE XIII

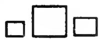
Percentage Passing the Selected Sentences at Each Age

Test and Direct'n Number	Revised Test Numb'r	Direction	Per cent passing at age— 6	7	8	9	10	11	12	13	14	15	16	17	Sup Ad.	Total	Credits
Easy Direcs.								(Based on 425 tested)									
16 E	1	Put a nose on this face.	36.2	64.3	78.3	100.0										90.8	1
10 E	2	Write 7 in the largest square.	45.4	60.5	78.3	94.5										89.2	1
5 E	3	Put a dot below this line.		50.5	62.5	88.5	97.5									84.0	1
11 E	4	Cross out the blackest letter in Texas.		42.0	62.6	82.6	97.5									83.3	1
12 E	5	Write g on the egg-shaped figure.		33.4	53.1	82.6	97.5	93.3								80.0	1
14 E	6	Put the sign = where it belongs.		39.6	50.0	70.6	87.5	90.0								78.4	1
7 E	7	Make a boy's name by adding one letter to Joh			25.0	58.8	92.5	96.7								70.8	1
19 E	8	Draw a line around the three dots.			50.0	64.9	85.0	76.7								66.6	1
Hard Direcs.								(Based on 844 tested)									
6 H	9	Now, if Tuesday comes after Monday, . . . here.			27.7	51.1	53.4	60.5	69.5	83.5						70.1	2
9 H	10	Show by a cross when the nights are longer . . .				37.8	28.8	56.3	51.7	61.5	69.2					57.0	2
8 H	11	Notice these two numbers· 3 5. If iron . . . here.					33.4	41.6	50.6	59.2	62.0					50.2	2
11 H	12	Do nothing here (5+7= . . .), unless . . . . line.						29.2	47.5	56.8	60.3	55.5				47.9	1+2
Harder Direcs.								(Based on 698 tested)									
8	13	If in the following words.								40.6	68.1	59.3	58.7	52.5	76.3	67.0	3
5	14	A list of words is given below. One of them is. .								46.9	47.4	57.1	58.7	55.7	83.0	66.5	3
6	15	You are looking at a clock face. . . . . . . at 3·45.								12.5	26.7	29.7	41.3	34.4	56.1	41.8	4

Name .... Age .... Grade ....

PINTNER-TOOPS REVISED DIRECTIONS TEST

Put a nose on this face: 

Write 7 in the largest square: 

Put a dot below this line: ————

Cross out the blackest letter in **TEXAS**

Write g on the egg-shaped figure: 

Put the sign = where it belongs: 3+2 5.

Make a boy's name by adding one letter to Joh

Draw a line around the three dots . . . .

Now if Tuesday comes after Monday, make two crosses here ...., but if not, make a circle here .... or else a square here ....

Show by a cross when the nights are longer: in summer?.......... in winter?..........

Notice these two numbers: 3, 5. If iron is heavier than water, write the larger number here ...., but if iron is lighter write the smaller number here ....

Do nothing here (5+7=......) unless you skipped the preceding question; but write the first letter of your first name and the last letter of your last name at the end of this line.

If in the following words e comes right after a more times than e comes just after i, then put a line under each word containing an e and an i, but if e comes just before a more times than right after i, then put a line under each word containing an a and an e:

receive feather teacher believe

A list of words is given below. One of them is needed to complete the thought in the following sentence: The roads became muddy when the snow..........

Do not put the missing word in the blank space left in the sentence, but put a cross below the word in the list which is next above the word needed in the sentence.

water

is

melted

snow

You are looking at a clock-face in a mirror. Place the hands of the clock at 3:45.

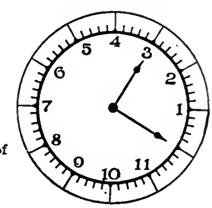


Figure 4.

Figure 4 shows the Revised Directions Test. The first eight directions are taken from the Easy Directions Test, the next four from the Hard Directions Test, and the last three are the three new harder directions which we have added.

The number of credits allotted to each direction is shown in the last column of Table XIII. This gives one point credit to the first eight directions and leaves the method of scoring these directions precisely as before. The details in regard to scoring have been explained above.

Two points credit are given to directions 9, 10, and 11, but *no partial credits are to be allowed.* These directions must be answered absolutely correctly in order to get a credit of 2; if not absolutely correct, *i. e.*, if partially correct, no credit whatever is allowed; a credit of 1 is never given to these directions. For direction 12 three credits are allowed, and these are divided into 1 and 2 credits, 1 for the first part of the direction and 2 for the second part. Each part must be absolutely correct in order to receive the partial credit of 1 or 2 as the case may be. The second part is always scored 2, for a correct response, or 0, for an incorrect response.

Directions 13 and 14 receive three points credit each, and no partial credits; the score is either 3 or 0. The correct response to direction 13 is the underlining of both "feather" and "teacher." The correct response to direction 14 is the placing of a cross (X or +) underneath the letters of the word, "is." No credit is allowed if the cross is on the word "melted;" it must be above this word. No credit is allowed if the word "melted" has been written in the incomplete sentence above, unless the word has been written in and crossed out again.

Direction 15 receives 4 points credit for an acceptable performance. An acceptable performance is interpreted as meaning that the minute hand on the diagram be drawn within one minute on either side of the figure, "9," and the hour hand shifted to the left anywhere up to the figure, "4;" *i. e.*, the hour hand must be shifted somewhat to the left in order to receive credit. No partial credit is allowed.

A method of procedure in giving the revised form of the directions test has not yet been standardized. In all probability the most feasible method will be to give as instructions to the subjects merely the words, "Fill this blank out according to directions. Do exactly what you are told to do on the paper," and to allow the subjects to

work at the test for a period of eight minutes. Each one is given as much time as he needs to complete the test, provided he does not exceed the time limit of eight minutes.

Norms for the Revised Directions Test

The directions, selected and arranged in order of difficulty, are useless without norms. Now, it seems quite possible that the total score made upon any direction depends somewhat upon its position with reference to other directions; so that in the rearranged form we cannot be sure of the sentences receiving the same percentages at each age as in the original. The only adequate thing to do, of course, is to restandardize the new test, and we are at present working upon this standardization.

Realizing the immediate need for norms, we have adopted the following procedure to secure norms which are, we believe fairly accurate and will serve our purposes until more accurate norms are determined: (1), all papers were regraded on the basis of the selected sentences and the revised credits (2), the norms thus obtained from each of the several tests for each age were added to secure the norms for the entire revised test. Since not the same but different groups (in part at least) of children were tested on the three tests, our procedure may here be questioned. Our justification and the assumption underlying this procedure is this: The *median* child, from whatever group, within proper limitations, he may be selected, should make the same score as any *median child of his age* on the combined list. We, therefore, are justified in crediting the median child (who determines our norm) at any age, say 16, with the median score made by the median child of age 16, on the other two tests without testing him on those tests. This is the same principle as is used in the Binet Scale which credits the child with all tests below his basal age. We do not feel justified in establishing percentile scores on this basis. We, therefore, present the norms (medians or 50-percentiles) for the revised test in Table XIV, based on 2366 individual measurements.

TABLE XIV

Tentative Norms for the Revised Directions Test

Test	Median Scores for Age— 6	7	8	9	10	11	12	13	14	15	16	17	Superior Adults
Easy Directions B	1	3	6	7	8	8	8	8	8	8	(8)	(8)	(8)
Hard Directions	(0)	(1)	2	3	3	4	5	7	7	7	6	7	9
Harder Directions	...	...	.	..	0	0	(0)	(0)	3	4	6	6	7
Revised Test	1	4	8	10	11	12	13	15	18	19	20	21	24

Summary

1. A standardization of the Woodworth and Wells Easy and Hard Directions Tests has been attempted. This standardization includes method of procedure, method of scoring and percentile norms for many ages.
2. An analysis of the relative difficulty of the directions in the two tests has been made, showing that many directions are of much the same degree of difficulty.
3. To secure one test covering a wider range of difficulty than any one of the original tests, a combination of the two tests as they are at present is not feasible.
4. A revised Directions Test has been constructed by choosing the most suitable directions from the original tests and by adding three other directions.
5. A method of scoring and tentative norms for the Revised Directions Test are offered.

A COMPARISON OF TWO TYPES OF LEARNING BY MEANS OF A SUBSTITUTION TEST

C. TRUMAN GRAY

University of Texas

It is the purpose of this article to report the results of an investigation in which two types of learning have been compared. Such a comparison is made from four different standpoints: (1) By means of learning curves, (2) by a compilation of errors, (3) by means of retention tests, and (4) by means of transfer tests. In order to get data upon these four points, the following general plan has been used: (1) the subjects were divided into two groups, (2) all the conditions under which the two groups worked were the same with one exception, (3) the one factor which varied with the two sections was the method of learning.

Material Learned

The material which was used is in the form of a substitution test. A brief description of this test was given by the author* some time ago. Later Baldwin† used the test upon delinquent girls, but no attempt was made by this writer to bring out the comparisons mentioned above. However, for the convenience of the reader it may be well to give a brief description of the test at this time.

Plate 1 shows one of the twelve sheets of the test. It will be noted that the letters of the alphabet are placed in a miscellaneous order at the top of the page and that each letter is accompanied by a number combination composed of two, three or four digits. In forming these combinations, only two different digits are used, namely, "1" and "2." In transcribing the material at the left of the page, it is understood that the digit "1" means a short horizontal line to the left of a given vertical line, while the digit "2" means a short horizontal line to the right of the vertical line. This phase of the test is made easy for the learner by the two expressions, "1 left" and "2 right," which are placed just beneath the alphabet. The vertical line is given in any case on the printed sheet. From the description which has been given, it will be seen that the letter H (12) is made with one stroke to the left and one to the right of the vertical while M (1221) takes the form left, right, right, left. Fur-

*Journal of Educational Psychology, Vol. IV., 293–297.

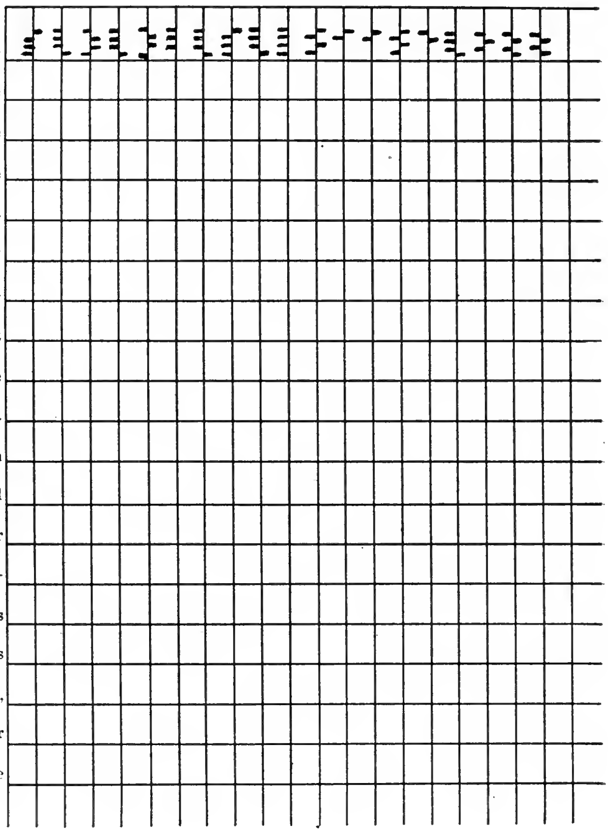
†Journal of Educational Psychology, Vol. IV., 317–332.

PLATE I

Z 2212 X 222 K 1211 G 1122 B 111 V 22 Q 212 J 122
P 211 W 221 F 1121 U 2122 L 1212 C 112 M 1221 Y 2211
D 1111 H 12 I 121 A 11 T 2121 E 1112 N 1222 S 2112 O 21 R 2111

1. **LEFT** 2. **RIGHT**

remembered, too, the ill-conditioned people whose disputes had been made part of his experience, the dogged children, cruel fathers, sullen husbands, angry wives, quarrelsome neighbors; and surely he did not err when he connected the two memories together. How many men and women go about pale-skinned and weak of limb because their physical health during infancy and childhood was not established by judicious management. It is just so, thought Frobel, with our minds. There would be



1761-1213-1M-4676

ther illustrations may be found in Plate 1, where the first line of printed material has been transcribed into the code. Since the vertical line is given on the printed page, it is clear that the subject has only to put down the short horizontal lines. In making these short lines the subject begins near the top of the vertical line and the order in which the horizontal lines are put down is determined by the order of the digits in the number combinations which accompany the letters.

Another important feature of the test is that the number combinations are made up according to a definite system. This plan shows itself clearly when the letters are arranged in their alphabetical order, as follows:

(1)	(2)	(3)	(4)
A 11	H 12	O 21	V 22
B 111	I 121	P 221	W 221
C 112	J 122	Q 212	X 222
D 1111	K 1211	R 2111	Y 2211
E 1112	L 1212	S 2112	Z 2212
F 1121	M 1221	T 2121	
G 1122	N 1222	U 2122	

From the above, it is clear that the alphabet is divided into four groups. In groups one, two, and three, there are seven letters each, while in group four, there are only five letters. In group one, each letter in the code begins with "11" and each combination after the first is formed by adding "1," "2," "11," "12," "21," or "22" in the order just given, to the "base " combination "11." In group two, each letter in the code begins with "12," and the combinations after the first are formed by making the same additions to the base "12" as were noted in group one. Groups three and four are very similar to groups one and two, and no further explanation of them is necessary. It is also important to note that the word "AHOV" gives the letters with which each of the four groups begins.

Subjects

The work with this test covers a period of four years. In this time, about four hundred subjects have taken part in the experiment. These persons were all university students who did the work as a required laboratory exercise under the direction of the author. As a check upon this laboratory work, forty persons have taken the test under the most careful experimental conditions. The results obtained from this group of forty did not differ in any essential way

from the earlier results, and it was not deemed necessary to carry the second type of work further. The entire group of forty did not take the test at the same time. They were divided into four sections on account of the greater ease with which it is possible to deal with a small number of subjects in such work. With the exception of certain check experiments, the results given in the pages following are taken entirely from this group of forty.

Methods of Learning

The test has been given in two different ways, which may be spoken of as Method I. and Method II. Those who worked according to Method I. had the printed sheet, as shown in Plate 1, before them at all times. This plan gave the subjects opportunity for reference to the code at the top of the page. The explanations and directions which it is necessary to give to the learner working according to such a method are very simple. It is sufficient for the subject to understand that each letter has standing for it certain digits, and that these digits are to be translated into short horizontal lines.

In the second method, the tops of the sheets were torn off across the perforated line (X) as indicated in Plate I. Those who worked according to this plan had the organization of the code explained to them. In making the explanation, the alphabet was placed on the blackboard and attention was called to the four groups of letters, to the "base" digits, to the additions made to the "base" digits, and to the word "AHOV." After the explanation had been given, and all questions answered, the subjects were given four minutes to study the plan or organization of the code. Before the subjects began this study period, emphasis was put on the fact that they were not to attempt to memorize the individual letters, but rather that attention should be given to the general plan underlying the code. Immediately after the study period, the experiment was begun.

All other phases of the methods were the same for each of the groups. The time spent was twenty minutes each day for twelve days. A record was kept of the five-minute periods. This was done by having a timekeeper call time at the end of each five-minute period. The subjects then made a mark at the point where they were working. In order that the experiment might have some motive, it was given after the learning curve had been discussed in class. As a result of this discussion, the general purpose of the experiment was understood, and most students showed a genuine interest in the work.

It is also of interest at this place, to compare the two types of learning. It is evident that the two methods are alike in that each requires such mental activities as memory, association, and the rational processes. The essential difference seems to lie in the fact that Method I. requires the learning of the letters as separate and distinct entities, while Method II. requires that the different letters be learned as parts of an intricate system which brings into relation all the letters of the alphabet. In other words, it may be said that Method II. requires the rational processes in a higher degree than does Method I.

It may be well to point out in this connection that Method I. parallels the so-called mechanical work of the school, while Method II. parallels the so-called rational work. Much has been written concerning these types of work, but very few suggestions have been made for a correct evaluation of them. Such experiments as the present one give certain indications, at least, as to their relative values, and it is hoped that a parallel experiment in one of the common branches can be carried out in the near future.

Results

Plate II shows a number of typical learning curves. The solid lines are curves made from the results of those working by Method I., while the dotted lines are the results of Method II. These curves have been selected for two purposes: (1) To show the extremes of the distribution of the different abilities brought out by the experiment, and (2) to make possible a comparison of curves which show about the same proficiency at the end of the training period.

In regard to (1) attention may be called to the fact that the upper limit of the distribution (Curves A and B) was attained by representatives of each of the methods, while the lower limit (Curve E) was passed by all those who used Method I. Five subjects, all of whom showed about the same ability as that in Curve E, were recorded. All of these used Method II. One of the poorest records made by Method I. is shown in Curve D.

The above results show clearly that Method II. makes for a wider distribution of ability than does Method I. If this fact is translated into terms of teaching activities, it probably means that the teacher who rationalizes her work will have a greater per cent. of failures than the teacher who does not so rationalize her work. If this point can be established, it will introduce a new factor into the distribution of grades. That is, if the grades of a teacher do not conform to

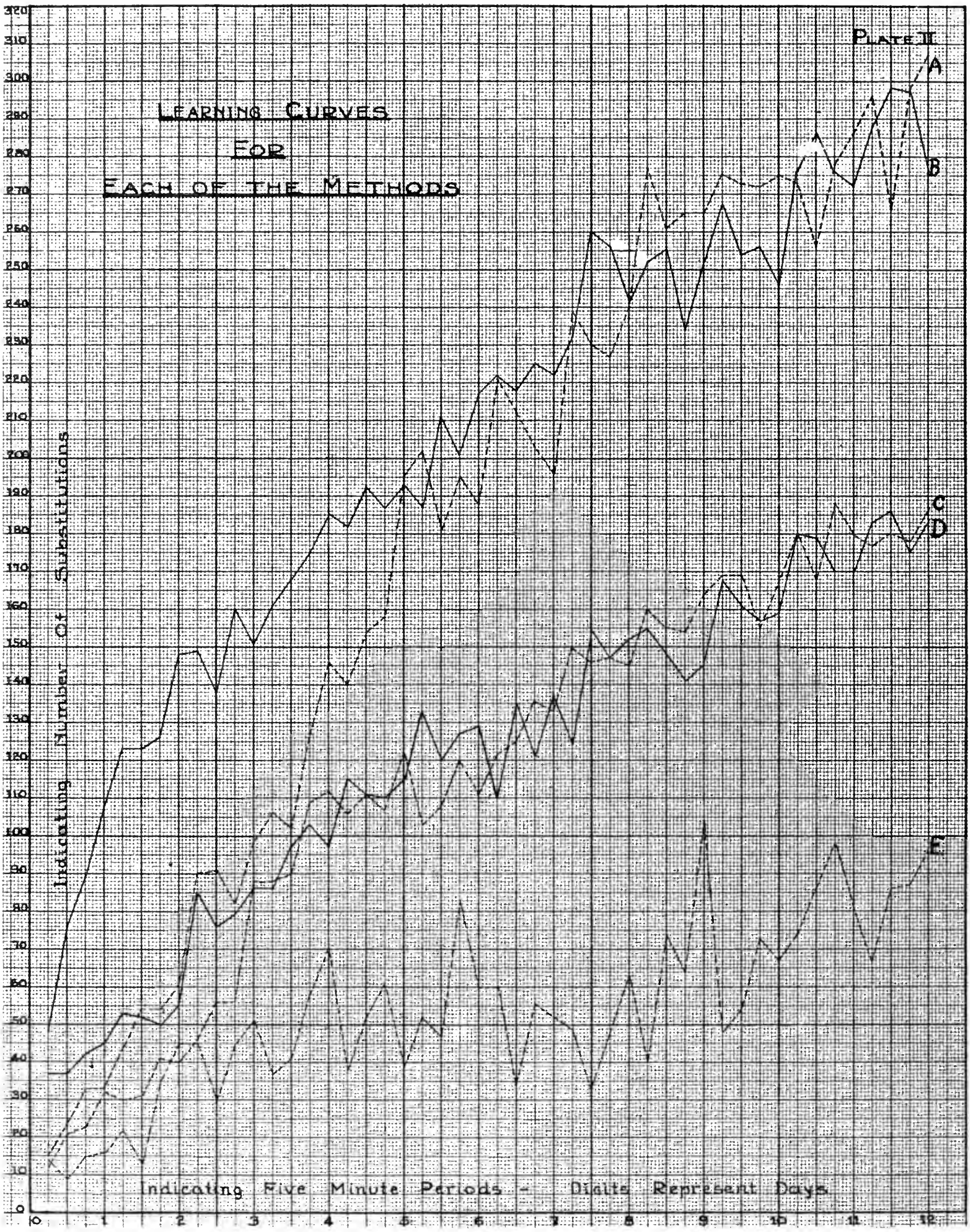


PLATE II
LEARNING CURVES
FOR
EACH OF THE METHODS
Indicating Number Of Substitutions
Indicating Five Minute Periods - Digits Represent Days
A
B
C
D
E
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
1
2
3
4
5
6
7
8
9
10
11
12

standards set for her, the results noted above give a basis for an explanation of the variation from the standard.

A comparison of Curves A and B shows that A begins lower than does B, and remains so until the fifth day, when the two curves cross. After this point they cross and recross until the end of the learning period. This same point is brought out in Curves C and D. Here the first crossing takes place on the third day after which the two curves cross and recross several times. This difference in the early parts of the curves is due, doubtless, to the fact that Method II. is more difficult in the beginning than is Method I.

A further study of Curves A and B shows a sudden rise in A on the third and fourth days, while a similar rise is to be noticed in Curve B on the first and second days. This same point is also shown in Curves C and D. This postponing of a rise in Curves A and C is evidently due to the fact that up to this period the time is consumed in a mastery of the organization of the code and the memorization of it. This means that Method II. requires a preliminary period of considerable length for adjustment before the progress of the learning shows in an objective way. In Method I. no such preliminary period is necessary, or if it is present, it is much shorter in duration. For certain subjects who worked by Method II. this preliminary period was lengthened and became a period of very great discouragement. During this period, not a few expressed themselves as desiring to discontinue the experiment. Their excuse was that they seemed unable to make progress.

In view of the large differences which were found to exist in the ability of the different subjects, it was thought worth while to compare certain individuals in each group, by means of other tests. Three different tests were given for this purpose. These were Thorndike's Association Test*, a test for the span of attention by means of Whipple's† revolving tachistiscope, and the tapping test.†† In selecting these tests, it was assumed that the ability to associate readily along with a wide span of attention and rapid motor reactions would make for efficiency in this experiment. These tests are so well understood that no further explanations are necessary except to say that the time allowed for study of the association test by the subject was two minutes, and that the tapping was done for thirty seconds. Table 1 gives results for the three individuals who did exceedingly well in the original experiment, three taken at random from those who showed mediocre ability, and three of the five who did most poorly in the work.

Educational Diagnosis, Science N. S. 37 No. 943, 133–142.

†*Manual of Mental and Physical Tests, Simpler Processes,* Page 265

††*Manual of Mental and Physical Tests, Simpler Processes,* Page 130.

TABLE 1.

Results in Auxiliary Tests

Association Test	Span of Attention	Tapping Test	
(1) 100	4 plus	6.7	
(2) 60	4	7.0	High (1)
(3) 100	4 plus	8.1	
Av. 86.6	4 plus	7.2	
(4) 100	4 minus	5.5	
(5) 70	4	6.8	Mediocre (2)
(6) 65	5	5.5	
Av. 78.3	4.3 minus	5.9	
(7) 80	4	8.4	
(8) 32	3 plus	5.8	Low (3)
(9) 60	4	5.6	
Av. 57.3	3.6 plus	6.6	

In considering these results, it will be noted that each of the individuals in group (1) stands high in all the tests, with the exception of No. 2, who is low in the association test. The low score in this one test may be accounted for by the fact that this individual failed to adjust herself readily to the experiment, and therefore did not do herself justice. In the middle group, it will be seen that each individual stands low in one or more of the tests. To illustrate, in-individual 4 has a span of attention, a little less than 4 and his rate of tapping is 5.5. Individual 6 is low in both the association test and the tapping test. In group (3), there are large individual variations in all the tests. It is rather interesting to note that the highest tapping rate is made by an individual in this group, and that the lowest score in both the association test and the span of attention test, is also made by an individual in this group.

If averages are considered, it is clear that the highest average in the association test is made by the upper group, that the next highest is made by the middle group, and that the lowest score is made by the last group. The correlation between this test and the results of the experiment can be accounted for by the fact that in learning the substitution test, ability to build up associations is used in a large degree until the test is completely mastered. In the

tapping and span of attention tests, the correlation is not so high as that just noted. This is probably due to the fact that in the substitution test, neither the entire span of attention nor the full motor ability is used at all times in the work.

Motor Test

The introspections of certain subjects who took part in the experiment indicated that a large part of the learning was motor in its nature, that is, more work could be done in five minutes near the end of the experiment because the subject was physically able to put down more marks to the right and to the left of a vertical line. While this point does not bear directly upon the comparisons to be made here, yet it throws light upon the learning involved in the experiment and for this reason data upon the question are given.

In order to get results upon this phase of the learning, the group of forty subjects mentioned before carried along with the regular work of the experiment, what may be spoken of as a motor test. This test consisted in making short horizontal lines on alternate sides of a vertical line for ten seconds each day before the regular experiment began. The period was made short in order that the practice element in the motor test might be as small as possible. Typical results of this test are given in Table 2.

TABLE 2.

Results for Motor Test

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1st day	28	27	27	27	24	16	27	33	18	24	26	21
2nd "	33	27	29	21	41	23	29	40	24	28	48	26
3rd "	32	27	34	26	41	23	36	58	25	34	49	26
4th "	23	35	39	29	45	24	33	64	27	36	45	34
5th "	35	36	47	31	48	27	37	58	30	40	44	42
6th "	33	31	48	32	53	36	34	56	30	40	52	39
7th "	36	34	49	29	49	31	37	52	28	40	54	43
8th "	36	28	50	40	57	31	40	65	26	42	54	40
9th "	34	31	46	40	53	34	46	55	29	47	56	39
10th "	40	38	54	39	59	32	49	66	30	50	39	47
11th "	36	38	56	42	54	40	48	65	32	50	51	47
12th "	37	44	56	49	58	40	55	63	32	48	64	51
A)	33.5	33.0	44.5	33.9	48.1	29.7	39.2	56.2	27.5	39.9	50.1	37.9
(B)	34	48	84	77	64	87	72	56	34	72	55	98
C)	204	313	269	214	236	317	606	776	403	525	490	816

Each column is the record of one individual. The row of figures indicated by the letter A at the bottom of Table II, gives the average number of strokes made for the twelve days. This varies from 27.5 to 50.1. Row B gives the per cent. of increase made in the motor test. This per cent. is based upon the average of the highest three and of the lowest three records. The numbers in the row indicate by the letter C give the per cent. of increase made in the substitution test. This increase is based upon the averages of the highest six and the lowest six records made in this test. If the per cent. of increase in the bottom row may be thought of as an increase due to both motor learning and intellectual learning, then the results in Row B may be thought of as due to motor learning alone. It is not good psychology to subtract one average from the other and say that the difference represents mental learning; yet it is safe to say that a large part of the learning is motor in its nature.

Column (1) is the record for the individual represented in Curve A, in Plate II. This individual complained that her hand was in her road. Column (2) is the record of a slow individual who complained that she could not learn the plan of Method II.

As a check upon this work, thirty subjects took the motor test once before they began the experiment and a second time after they had finished. These results show the increase in motor ability due to the practice of the experiment alone, without the practice of the motor test being introduced. Per cents of increase for fifteen persons taken at random from this group are as follows: 44, 24, 30, 14, 26, 47, 55, 36, 21, 34, 52, 46, 63, 142, 178. The average of the above is fifty-four per cent. This is eleven points lower than the average of the results in Row B of Table 2. This shows that a part of the increase to be found in Table 2 is due to the daily practice period, of ten seconds, yet the results in Table 2 are not to be considered as invalidated. In other words, the conclusion holds that a considerable portion of the learning in the substitution test is motor in its nature.

A Comparison of Errors Made in the Two Methods

It was thought worth while to go over a number of the records made in the substitution test in order to check the errors made. The errors found consisted in the omission of letters and in mistakes in putting down the short horizontal marks. In the beginning of the experiment, no emphasis was put upon the fact that the records would be checked for errors. Table 3 gives typical results

TABLE 3.

Showing Errors Made in Each Method

	Method I.					Method II.				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1st day	2	1	2	6	4	0	16	11	13	2
2nd "	3	8	0	0	5	0	36	4	8	7
3rd "	5	2	2	3	0	3	6	0	7	2
4th "	2	1	7	2	4	0	1	4	11	68
5th "	5	3	2	0	6	2	3	2	1	2
6th "	3	0	2	3	0	4	5	4	2	16
7th "	4	16	3	1	3	5	2	5	1	2
8th "	4	2	2	2	6	7	1	6	0	5
9th "	1	4	6	0	3	9	0	0	0	4
10th "	7	1	1	3	2	2	7	0	4	11
11th "	7	2	1	3	1	2	0	1	0	5
12th "	4	1	1	8	2	5	1	6	9	7
Total Errors	47	40	29	31	36	39	78	43	47	120

in this work. The columns numbered (1), (2), etc. show the number of errors made by different subjects on the twelve days of the test.

The general conclusion to be reached here is that in the earlier periods of the training there are more errors made in Method II. than in Method I. A rather marked exception to this is the record in column 6. In the later part of the experiment there is very little difference in the number of errors shown in the two methods. In other words, the two methods give about equal results in accuracy at the end of the training period.

The large number of errors made in Method II. at the beginning of the experiment is due to the fact that this type of work is entirely mental, while in Method I. there is opportunity for reference to the code at the top of the test sheet. It is apparent that as the subject becomes more and more familiar with the code in Method II., the errors will be reduced and at the end of the training period there is little, if any, reason why one method should afford more errors than the other.

Retention Tests

Attention has been called to the fact that those working by Method II. met certain rather serious difficulties at the beginning of the experiment.

It is also true that Method II. requires much greater mental effort than Method I. This was clearly shown by the wrinkled brows, the complaints of headache, and remarks concerning the close concentration required by the work. This does not imply that those who worked by Method I. did not put their best efforts into the experience, but rather that they were not put to the same severe mental strain as those who worked by Method II.

The question may now be raised as to whether there is any compensation for this period of discouragement and the extra effort required by Method II. This point was dealt with by means of a retention test. The first test used in this phase of the work consisted in having the subjects write the alphabet in the code as quickly as possible two weeks after the experiment was completed. This test was found unsatisfactory because there were always letters which those working by Method I. had not memorized. These letters could not be put down in the retention test; yet, in carrying out the experiment, there seemed to be a motor habit of the eye, which enabled the subject to locate such letters very quickly in the alphabet at the top of the page.

This type of test was also found unsatisfactory because no time limit could be established. If the time was made short, very few would finish the alphabet, and if the time was made longer, then all those who had worked by Method II. would finish the entire alphabet on account of the relations existing among the various letters.

A second type of retention test consisted in having all subjects work upon the regular test for ten minutes. The first time this test was given, the tops of the sheets were torn off for all subjects, but this was found to be unsatisfactory again on account of certain letters not being memorized, as was explained above.

The test finally used, consisted in having the subjects do a certain part of the test in the same way that the regular experiment was carried out. The subjects did not know that such a test was to be given, so that there was no voluntary effort at retention. The time given to the test was ten minutes, and the record was kept in five-minute periods, as in the regular experiment. The period which elapsed between the end of the regular experiment and the retention test was two weeks. Table 4 gives typical results in this test:

Column 1 under both methods, gives the average for the highest six records made in the regular experiment as mentioned previously, while column 2 shows the average for the two five-minute periods of

TABLE 4.

Results for Retenion Tests

Method I. (1)	Method I. (2)	Method I. (3)	Method II. (1)	Method II. (2)	Method II. (3)
175.5	136.0	22	189.0	142.5	24
140.5	64.5	54	142.5	114.0	20
164.3	68.0	58	198.5	172.0	13
166.5	70.0	58	313.0	284.5	9
147.2	75.0	49	111.5	31.5	44
172.5	112.5	34	75.5	61.5	18
123.0	50.5	59	116.5	83.5	28
148.0	105.0	29	151.6	113.5	25
152.8	132.5	13	136.0	84.5	38
217.5	184.0	15	109.0	66.5	39
211.0	171.5	18	140.0	114.5	18
	Av.	37.1		Av.	25.0

the retention test. Column 3 gives the loss in per cent. for each of the methods. The average of the different losses for Method I, is 37.1, while the average for Method II. is 25.0. Thus Method II. has a distinct advantage over Method I. in this part of the work. This is probably due to the fact that the extra effort required by Method II. brings the learning to a higher level than does Method I.

Transfer Tests

Another method of comparing the two types of learning consists in determining the amount of transfer from the training given in the substitution test here used, to other tests of the same type. Those who worked by Method I. did ten minutes of work upon Dearborn's* test before they began the twelve days' training, while those who worked Method II. worked for ten minutes upon a test used by Goddard.†

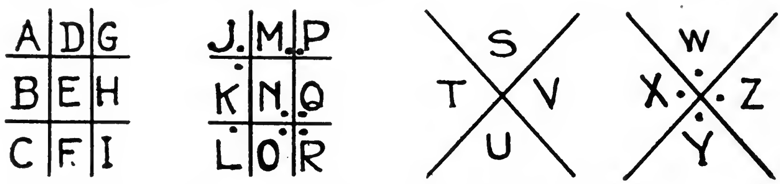
The Dearborn test consists in transcribing letters into digits. Each letter of the alphabet along with the number which stands for it is placed in a circle: thus

*Psy. Rev. Mon. Sup. Vol. 13.

†Holmes. *School Organization and the Individual Child,* Page 108.

The circles with the letters and numbers included are grouped together in a miscellaneous order at the top of the test sheet. These headings were clipped from the original sheets and pasted on the sheets of the substitution test used in the present experiment. This was done in order that all the material to be translated might be homogeneous.

The Goddard Test is as follows:



According to this code the word "horse" appears thus:

This code was made out upon slips of paper and used in the manner just indicated for the Dearborn Test.

After the training period, the Dearborn and Goddard tests were given a second time in order to determine the increase. Tables 4a and 5 give typical results in this work.

TABLE 4a.

Increase Shown in Dearborn Test After Practice in Method I.

B. P.	A. P.	Per Cent of Increase
104.0	118.0	13
107.0	115.0	7
103.0	128.0	24
154.0	171.0	11
88.0	111.5	26
90.0	116.0	29
87.5	113.5	52
101.5	133.5	31
45.5	82.5	81
72.0	104.5	45
		Av. 31.9

In the above tables, the column headed B. P. shows the number of substitutions made in the test before practice, the column headed A. P. gives the number of substitutions after practice, and the third column gives the per cent. of increase.

If the averages are considered, the training in Method II. is seen to result in greater transfer than Method I.

TABLE 5.

Increase Shown in Goddard Tests after Practice in Method II.

B. P.	A. P.	Per Cent of Increase
59.0	93.5	58.5
52.5	87.5	66.6
46.5	56.0	20.4
39.0	43.5	11.5
30.5	43.5	42.6
36.5	62.0	69.0
37.0	72.0	94.6
33.5	51.0	52.2
28.5	43.0	50.9
	Av.	51.9

TABLE 6.

Showing Increase in Dearborn Test without Practice after an Interval of Twelve Days

B. I.	A. I.	Increase in Per Cent.
118.0	152.0	28.2
92.5	118.5	25.9
86.5	96.5	11.5
95.0	125.0	31.5
94.5	114.5	21.1
87.0	109.0	25.2
111.0	135.0	21.6
94.5	101.5	7.4
77.0	99.0	28.5
47.5	69.5	46.3
	Av.	24.7

TABLE 7.

Showing Increase in Goddard Test without Practice after an Interval of Twelve Days

B. I.	A. I.	Increase in Per Cent.
76.0	84.0	10.5
100.0	127.0	27.0
54.0	81.0	61.1
99.0	135.0	39.1
83.0	87.0	4.8
20.5	35.5	73.1
28.5	59.5	75.0
57.0	64.0	12.1
35.0	46.0	31.4
47.0	55.0	17.0
	Av.	35.1

A part of the increase indicated in Tables 4a and 5 would have taken place if no practice had been given. In order to check up on this, control sections took the Dearborn and Goddard tests before and after an interval of twelve days with no practice. Results of this work are given in Tables 6 and 7.

The column indicated by B. I. gives the number of substitutions before the interval, and the one indicated by A. I. gives the number of substitutions after the interval, and the third column shows the increase in per cent. The average increase from the Dearborn test is 24.7 per cent., and that from the Goddard test is 35.1 per cent. The greater increase in the Goddard test is probably due to two things: (1) Since there are just two periods of work, and since those using this test began lower on the average, they have a possibility of going higher in the second period of work, and (2) the Goddard test requires greater concentration of mind, and it is probably true that this concentration gives better results.

If the average of the results of the Dearborn test in Table 6 is taken form the average in Table 4a (31.5–24.7), the result is 7.8 per cent, and if the average of the Goddard test in Table 7 is taken from the average in Table 5 (51.0—35.1), the result is 16.8 per cent., which gives the training by Method II. a decided advantage.

It is interesting to attempt an explanation of the greater transfer in Method II by means of the different theories of transfer which have been proposed. It is difficult to see how Thorndike's* theory of identical elements can explain the difference, because it would appear that there are as many common elements in one case as in the other. It is also difficult to see how Bagley's** theory of ideals can apply to either of these methods because ideals have little, if anything, to do with such work.

On the other hand, Angell's† theory of attention gives some basis for an explanation. It has been pointed out that Method II. requires greater mental effort than Method I., and it seems entirely possible that this training in attention may account for the greater transfer in Method II. It is also true that Method II. requires the higher mental processes to a greater degree than does Method I, and it may be that certain generalizations are required in Method II. which enhance the transfer in this method. This explanation would correspond to Judd's†† theory of transfer.

*Pys. R. 8, 247–261; 348–395; 553–564.

***Educative Process.* Chap. 13, page 208.

†Ed. Rev. June, 1908.

††Ed. Rev. June, 1908.

THE ECONOMIC STANDING OF PARENTS AND THE INTELLIGENCE OF THEIR CHILDREN

ARTHUR W. KORNHAUSER

University of Pittsburgh

I. Purpose

While the vanguard of the students of heredity and of educational psychology have gone forward with their inquiries into the relative importance of nature and nurture in determining the mental life of children, a very considerable body of social workers, teachers, and students, have failed to acknowledge the fundamental premise of these studies and hence, even at the present time, the problem persists as to whether more intelligent and wealthier parents do have, in the long run, more intelligent children. Some who maintain the affirmative have ascribed the superior intelligence of the offspring to heredity; some have believed early environment all-powerful; but a third group, less clearly recognized, has tended to view dubiously the entire proposition that children do vary with the standing of their parents. This group has emphasized the instances wherein great leadership or genius has apparently arisen from obscure parentage amidst abject poverty; they discredit family trees on the one hand and discount the effect of economic surroundings on the other hand.

The purpose of the following inquiry was simply to attempt, by a brief survey, to find whether any clear indication would appear of a correlation between intelligence of children, measured crudely by school advancement, and the economic standing of the parents, measured still more crudely. It was not proposed to make a comparative study of the influence of heredity and environment, but to ascertain what light might be thrown upon the position held by the third group mentioned above—the position, namely, that the economic attainment of parents (and hence also their degree of intelligence) is not definitely associated with the "brightness" of their children. Our question then is: Do poorer parents on the whole have less advanced children; do wealthier parents have more advanced children?

*A study made in connection with a research course under Prof. Roswell H. Johnson at the University of Pittsburgh.

II. Method

As an index for the measurement of the intelligence of children their school advancement was employed, an index well justified by the close association which has been repeatedly demonstrated between it and a ranking based on individual mental tests. Only three class divisions were used which we shall call Retarded, Normal, and Advanced. Normal was used to signify that the individual's actual school grade (to the half year) is the same as, or within one half year of, his theoretically normal school grade; Retarded to signify his actual grade to be one year or more under the theoretically normal; and Advanced one year or more above the theoretically normal. That is:

Retarded:	(Age to nearest $\frac{1}{2}$ year—6) $\pm \frac{1}{2} >$ Actual School Grade
Normal:	(Age to nearest $\frac{1}{2}$ year—6) $\pm \frac{1}{2} =$ Actual School Grade
Advanced:	(Age to nearest $\frac{1}{2}$ year—6) $\pm \frac{1}{2} <$ Actual School Grade

In reckoning actual school grade, grade IA = 1 year, grade 1B = $\frac{1}{2}$ year, etc. For example, if X was born in January 1906 and his grade consulted in March 1917 was found to be 6B (equal to $5\frac{1}{2}$), then $(11-6) \pm \frac{1}{2} = 5\frac{1}{2}$. Hence X is Normal.

It was found that this method of ranking gave a more nearly normal distribution than any other tried and it also seemed most frequently justified in individual cases as actually indicating whether the child had failed of promotion at least once, in the case of Retarded, or had succeeded in advancing by double promotion, in the case of Advancement.

As an index for the measurement of economic standing it was desired to find a means of roughly dividing the parents into two groups—more wealthy and less wealthy. The possession of telephones was hit upon as a simple, readily accessible, and withal fairly trustworthy, indication of economic and social status. Both theoretically and empirically it was found that the possession of a telephone is rather reliable evidence that the parents have succeeded in attaining considerable economic independence and hence also that they are of comparatively high grade natural ability. Similarly absence of telephones points to a much less advanced social position and hence presumably to a comparatively lower grade of heredity constitution. In those numerous cases where the economic margin of surplus is such that telephones might or might not be had, in most instances, it would seem, the determining factor for such possession would be cultural attainment or comparative im-

portance of work, thus bringing these marginal cases into their proper significance in our inquiry. The use of this index is further justified by the actual results obtained as to the distribution of telephones, as will appear under our results.

The material to which our standards were now applied was a group of one thousand school children of Pittsburgh, distributed in five public schools. The five schools were chosen to be as representative as possible of different social settings: Grant and Ralston are in very poor districts with many laborers, day workers, street venders, etc; Bellefield and Shakespeare are in districts considerably more prosperous though by no means wealthy, populated to a considerable extent by artisans, small shop-keepers, boarding-house proprietors, etc.; Linden is one of the most wealthy districts of the city, drawing mainly from families in very comfortable circumstances or better. The significance of these differences both as to advancement of children and possession of telephones will be evident below.

The necessary data were obtained from the Permanent Record Cards of the pupils in each school together with reference in each case to the two city telephone directories. No children whose parents' present address was not definitely ascertainable were taken; all pupils entering in the current year as well as all kindergarten pupils were omitted; and in general precautions were observed to make the results as trustworthy as possible.

III. Results

TABLE I.

Distribution of 1000 School Children as Retarded, Normal, and Advanced

	Individuals				Per Cent.		
School	Retarded	Normal	Advanced	Total	Retarded	Normal	Advanced
Grant	47	49	8	104	45.2%	47.1%	7.7%
Ralston	25	38	5	68	36.7	55.9	7.4
Bellefield	122	209	84	415	29.4	50.4	20.2
Shakespeare	78	140	53	271	28.8	51.7	19.6
Linden	18	89	35	142	12.7	62.7	24.6
Totals	290	525	185	1000	29.0	52.5	18.5

In Table I are presented the general class distributions of the different schools and the totals both by count of individuals and by per cents. The percentage distribution in the schools is in striking agreement with what might be expected from the economic con-

dition of the respective districts. Grant and Ralston show a very large proportion of Retarded with an almost negligible number of Advanced; Bellefield and Shakespeare have the most nearly normal distribution of Retarded and Advanced; Linden shows the opposite tendency from the first two schools, a very small proportion of Retarded and with a comparatively large percentage of Advanced pupils; Linden shows the opposite tendency from the first two schools, a very small proportion of Retarded with a comparatively large percentage of Advanced pupils. This data in itself gives some indication of the marked association between economic status and school advancement and undoubtedly would be much more striking if the different schools had a system of uniform grading, for there can be no question that there is a tendency for the general lower ability in the poor school to be compensated by a general lower standard of grading and vice versa in the wealthier school.

TABLE II.

Distribution of Telephones among the Same 1000 Children

	Number of each class having telephones				Per cent of each class having telephones			
School	Retarded	Normal	Advanced	Total	Retarded	Normal	Advanced	Total
Grant	3	3	3	9	6.3%	6.1%	37.5%	8.6%
Ralston	3	2	1	6	12.0	5.3	20.0	8.8
Bellefield	28	64	38	130	22.9	30.6	45.2	31.3
Shakespeare	12	42	23	77	15.4	30.0	43.4	28.4
Linden	10	57	27	94	55.5	64.0	77.1	66.2
Totals	56	168	92	316				

Table II presents the distribution of telephones among the 1000 individual studied, again according to school and also according to the class (Retarded, Normal, Advanced) of the pupils. The total percentages of telephones show how definitely the districts are ty pified by this index, the range being from 8 per cent. to 9 per cent. in Grant and Ralston to 66.2 per cent. in Linden, with the other two schools set well off in the interval at about 30 per cent. The general tendency in the distribution of telephones among the classes is also clear, the last three schools presenting marked differences between each two classes. In the case of Grant and Ralston the difference

between the percentages of Advanced and Retarded having telephones is equally definite. The slight discrepancy in these schools between the proportions of Retarded and Normal possessing telephones is readily explainable by reason of the very few cases here involved (the fewest occurring throughout the tables).

TABLE III.

Per Cent of Class Totals Having Telephones

Retarded	Normal	Advanced	Totals
19.3%	32.0%	49.7%	31.6%

In Table III are summed up the few most significant percentages of the foregoing data. It is found that in the entire 1000 cases telephones are possessed by 31.6 per cent.; in the entire class of pupils in their Normal grade 32.0 per cent. have telephones, a percentage almost identical with that of the totals. In contrast to this similarity is the variation shown in the per cent. of Retarded children having telephones, 19.3 per cent., and of Advanced children having telephones, 49.7 per cent. Evidently a real association does exist in this case between Retarded Grade and Absence of Telephones and between Advanced Grade and Presence of Telephones.

In order to combine this association into a single approximate coefficient the four classes Advanced and Telephones, Advanced and Not Telephones, Retarded and Telephones, Retarded and Not Telephones, were utilized in the simple association formula of Yule, resulting as seen in Table IV in a positive association coefficient of .61.

TABLE IV.

Coefficient of Association of School Standing and Possession of Telephones=+.61

$$\text{Formula, } Q=\frac{N.\ \delta.}{(AB)\ (ab)+(Ab)\ (aB)}$$

(Cf. Yule: "Introduction to the Theory of Statistics", p. 38)

IV. Conclusions

In a few words we may summarize these results:

1. The proportion of families possessing telephones is markedly greater as the district studied is wealthier.

2. The proportion of school children who are advanced is markedly greater as the district studied is wealthier (despite the compensatory tendency of the grading).

3. In every district studied the tendency is for the percentage of Advanced pupils having telephones to be greater than that for the Normal and still greater than that for the Retarded.

4. The percentages of the respective class totals having telephones are approximately 19 per cent. of the Retarded, 32 per cent. of the Normal, 50 per cent. of the Advanced.

5. The coefficient of association between advanced grade and telephones and retarded grade and not telephones is +.61.

If, now, we can accept the indices used to actually represent parental attainment and intelligence of offspring respectively, and there appears ample justification for doing so in a general way, we may indicate the following few conclusions and comments:

6. The percentage of Advanced children coming from homes better stiuated economically is clearly very considerably greater than the percentage coming from homes that are poorer, and this applies as between the rich and middle class as well as between the extremes. The converse is of course true of Retarded children.

The predominant cause of these relations may be hereditary or environmental or both, that is,—

7. Parents having telephones are inherently of good enough stock to have succeeded and hence naturally their offspring are also of superior native ability.

8. Parents having telephones presumably also have homes in which the physical and educational environment are favorable to the children's mental life and school success.

9. The most reasonable view would seem that both the preceding factors are involved in varying proportions together with numerous minor causes. Whatever be the ultimate explanation, however, the data point to a very real association between parental standing and intelligence of offspring, worthy of much more careful analysis and evidently of no inconsiderable social significance.

The Journal of Educational Psychology
INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD.

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

THE JOURNAL is issuing a series of Monograph Supplements for the publicaiion of more extensive reseaiches in educational psychology. For details prospective contributors may apply to H. E. Buchholz, 10 East Center St., Baltimore, Md.

The price of the JOURNAL is $3.00 per year of ten monthly issues, if paid in advance. Single issues are 50 cents each. Back volumes are $3.50 in original wrappers and $4.75 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME IX MARCH, 1918 NUMBER 3

EDITORIAL

A New Humanism Needed

The classicists are growing desperate. They feel that they must put forth every effort "to defend the citadel of culture against the onrushing hoard of rude barbarians." Their monopoly of secondary education was broken a generation ago when the natural sciences fought and won their way to recognition in the course of study. But the sciences were parvenu, their teaching was poorly organized, and the tradition of culture was against them. Thus the classics held on. Greek, it is true, practically disappeared, but Latin took an upward turn, and experienced what might justly be termed a revival. Now comes the experimentalist, and with clear, unfaltering eye and steady, relentless tone, he demands of each subject the justification for its existence. Well may the Professor Shoreys, the Principal Stearns, the Dean Wests and other Latinists rush to the defense. They realize that behind President Eliot and Dr. Flexner is the grow-

ing host of scientific students of education in this country, that behind these are the rank and file of public school men, and that behind these are the American people, asking what does the study of Latin contribute to the well-being of the coming generation.

In the defense of their cause the Latinists fall back upon medieval nomenclature, and cry out that this is an "Assault on Humanism." They ignore the fact that the term "Humanism" was applied to the effort to get away from the pedantic formalism of scholastic hair-splitting, and to return to the rich, free, many-sided intellectual development of the Greeks. They ignore the fact that this movement speedily relapsed into a linguistic and grammatical formalism that was more barren than the one which preceeded it. They ignore the fact that this narrow linguistic "Ciceronianism" has persisted in classical studies from the time of Erasmus to the present day. The technical, linguistic training afforded by the Latin courses in our high schools and colleges has not even as much right to the term "Humanism" as that given by courses in stenography, cooking or carpentry. For the latter courses are less abstract, less formal, less remote from present day human interests.

The great need of our elementary and secondary education today is a reorganization on the basis of a "New Humanism." We need to make a systematic inquiry to determine what lines of contemporary human thought, human activities, human striving are most significant for and can be best utilized in educational procedure. Far from being inimical to a genuine Humanism, the experimental movementsets itself the task of freeing education from the shackles of a traditional, grammatical Latinism, and of really humanizing it. Have the literature and history of Greece and Rome anything to offer us that will satisfy present human needs? We believe that they have. Then let us drive directly at that which is humanly significant in those records, and not waste years of time and effort on linguistic difficulties, which only a small percentage of pupils surmount, and which effectively blind the rest to any humanizing influences that the student of classic civilizations might discover.

This is an age of momentous happenings. Great changes are taking place around us. Great questions are being fought out. Great social movements are under way. When education demands so much of us in such a short time, is it wise, is it sane, is it humanizing to spend from four to six years on the study of a language which we shall never use when school days are over? J. C. B.

NOTES AND NEWS

A department of vital economics has been established at the University of Rochester by the bequest of Lewis P. Ross, who directed that the income of his residuary estate of more than $800, 000 should be used "to the end that human life may be prolonged with increased health and happiness."

The Connecticut State Department of Health points out in a recent bulletin that it is both safer and cheaper to keep the schools open in time of an epidemic of contagious disease, and establish a rigid quarantine and careful examination of the pupils, than it is to close the schools and throw the pupils together on the street with little or no protection.

A Ministry of Public Health and Social Welfare has been constituted for Austria with Dr. Baernreither as the first minister. It is to supervise the care of war invalids, to combat war diseases, and to centralize pre-existing, unco-ordinated departments of public health and sociology. It is to have the care also of the dependents of fallen soldiers, infant welfare, housing and insurance.—*Science.*

In a course of public lectures on "Animal Life and Human Progress" recently given at King's College, London, we note the following titles: "Man's Account with the Lower Animals," by A. Dendy; "Some Educational and Moral Aspects of Zoology," by G. C. Bourne; "Man and the Web of Life," by J. Arthur Thompson; "The Origin of Man," by F. Wood Jones; and "The Future of the Science of Breeding," by R. T. Punnett.

At the meeting of the New England Association of Colleges and Secondary Schools held at Boston, March 9, the subject for discussion was "The Modern School." The speakers were Charles W. Eliot, Harvard University, Paul Shorey, University of Chicago, Alfred E. Stearns, Phillips Andover Academy, and Otis W. Caldwell, the Lincoln School.

The Bureau of Educational Measurements and Standards, of the Kansas State Normal School, announces a Standardized Silent

Reading Test. This has been constructed on the lines of the old Kansas silent reading tests, but the exercises have been selected from school readers and other books which children read. Thus the criticism made of the old tests, that they are largely mathematical puzzles, is forestalled, and a reading situation is furnished that is more nearly normal. Dr. Monroe has promised an account of the standardization of this test for an early numberof the JOURNAL.

At the February meeting of the New York Branch of the American Psychological Association the following papers were presented: The Influence of Practice on Correlation of Abilities, by Georgina Strickland; A Tentative Formulation of a Psychology of Play, by Clara F. Chassel; Families of American Men of Science, by Dean R. Brimhall; A Note on a Mathematical Prodigy, by Lorle I. Stecher; and Redintegrative Mechanisms in the Psychoneuroses, by H. L. Hollingworth.

The division of medicine of the National Research Council has perfected an organization of committees to carry on war research work and to keep in close touch with research progress and problems in Europe. The Psychology Committee, Major Robert M. Yerkes, Chairman, includes the following subcommittees: (a) Methods for the Psychological Examination of Recruits, Robert M. Yerkes, Chairman; (b) Tests of Special Skill, Edward L. Thorndike, Chairman; (c) Problems of Aviation, including the Examination of Aviation Recruits, Edward L. Thorndike, Chairman; (d) Incapacity, Re-education and Vocational Training, Shepherd I. Franz, Chairman; (e) Visual Problems, Raymond Dodge, Chairman.

Dr. Bird T. Baldwin, director of the child welfare research station, University of Iowa, has been commissioned major in the sanitary corps of the United States Army in connection with the psychological examination of recruits. His place in the University of Iowa will be filled temporarily by Professor E. E. Faris.

Professor M. E. Haggerty, of the University of Minnesota, has received a commission as major in the sanitary corps of the army. Major Haggerty will have charge of psychological work in connection with special hospitals and the re-education of wounded and disabled soldiers.

The following psychologists have received commissions in the sanitary corps for psychological testing in the army: Albert T. Poffenberger, instructor in psychology, Columbia University, captain; Garry C. Myers, department of psychology, Brooklyn Training School for Teachers, lieutenant; Schachne Isaacs, instructor in psychology, University of Cincinnati, lieutenant (psychological research in high altitude aviation); Roberts B. Owen, instructor in philosophy, Columbia University, lieutenant.

Dr. Henry Maudsley, the distinguished British alienist and psychologist, has died at the age of eighty-three years.—*Journal of Philosophy, Psychology and Scientific Methods.*

Dr. Henry H. Goddard, for ten years head of the research department of the Vineland Training School, has been appointed head of the Bureau of Juvenile Research of the State of Ohio.

Dr. Edwin Bissell Holt, assistant professor of psychology at Harvard University, has tendered his resignation to take effect September 1, 1918.

Mr. John F. Arundel, principal of the Lincoln School, Cincinnati, has been appointed to the newly created post of director of vocational education in Cincinnati.—*School and Society.*

Frederick W. Steacy (Ph.D., Teachers College, Columbia University) has been appointed lecturer in education at McDonald College, the teachers' training department of McGill University, Montreal, Canada.

PUBLICATIONS RECEIVED

FRANCES R. ANGUS. *Fundamentals of French.* New York: Henry Holt and Company, 1916. Pp. xv, 280.

This admirable introductory French book offers a combination of the direct and the grammar methods. There are many stories illustrating difficult grammatical constructions, there are model letters giving drill on practical forms, and there is ample provision for conversational material.

WALTER S. ATHEARN. *Religious Education and American Democracy.* Boston: The Pilgrim Press, 1917. Pp. xi, 394. $1.50.

A program for religious education similar to and parallel with the secular public school education. The separation of church and state is enthusiastically defended, but the church needs to offer a more adequate and systematic religious education. The book shows sound scholarship and sympathetic vision.

ALBERT F. BLAISDEL AND FRANCIS K. BALL. *American History for Little Folks.* Boston: Little, Brown and Company, 1917. Pp. x, 138. $.75.

A series of charming stories about dramatic incidents in the history of our country. The book is intended for children of the third and fourth grades.

EMMA MILLER BOLENIUS. *Everyday English Composition.* New York: American Book Company, 1917. Pp. xii, 340.

Direct, vivid, practical, crowded full of valuable points. An interesting feature is the abundance of tiny illustrative drawings in the text. The chapter on reporting, special articles, and editorials will be especially useful in connection with the school paper.

BURTIS BURR BREESE. *Psychology.* New York: Charles Scribner's Sons, 1917. Pp. x, 482.

This is distinctly a sensationalistic rather than a behavioristic psychology. The neural substrate of mental activity is richly illustrated with excellent diagrams, and there are well-selected figures of the sense organs and of visual illusions. The motor aspect of consciousness is treated very briefly, the subject of habit being dismissed in a few sentences in the eighteenth chapter. The text is avowedly eclectic, and aims to present the generally accepted facts of modern psychology. This aim is attained with a high degree of success.

H. A. BROWN. *The Measurement of Ability to Read.* Concord, N. H.: Bureau of Research, State Department of Public Instruction, 1916. Pp. 57.

The fact that a second edition of this study of the measurement of reading has been called for is an indication of the rapidly growing interest in the subject. It is to be hoped that the author will give us the results of the further application of his test.

STEPHEN SHELDON COLVIN. *An Introduction to High School Teaching.* New York: The Macmillan Company, 1917. Pp. xxi, 451. $1.60.

This book is not a treatise on high school organization and administration, nor does it concern itself greatly with the specific subjects of the high school course of study. It is rather a discussion of general methods in secondary teaching. There is an introductory consideration of the nature of secondary education, the high school pupil, and the high school teacher, followed by three chapters on discipline, indirect and direct control, and the function of punishment. The remaining chapters develop such aspects of method as elimination of waste, testing knowledge, conduct of drill, oral and book instruction, illustration and demonstration, stimulating pupils to think, inductive and deductive development lessons, the use of the question, lesson plans, and supervised study. When principals and superintendents begin to demand of their high school teachers the study of such books as this, the teaching in our high schools will show a vast improvement.

KARY C. DAVIS. *Productive Plant Husbandry.* Philadelphia: J. B. Lippincott Company, 1917. Pp. xvi, 462.

With such stimulating and scholarly text-books the courses in high school agriculture should rise to a high plane of popularity and value. The present book deals with plant selection and breeding, soils, field crops, gardening, fruit growing, forestry, control of insects and plant diseases, and the outlines of farm management. There are over three hundred splendid illustrations.

GEORGE R. DAVIES. *Social Environment.* Chicago: A. C. McClurg and Company, 1917. Pp. 149. Fifty cents.

The author accepts the biological and evolutionary point of view as the basis of sociological theory, but contends that the higher spiritual values, which with advancing civilization become more and more clearly conscious social ideals, have failed to receive due weight in the explanation of social progress. Scientific social organization will take account of ethical idealism as the most significant dynamic factor, and will strive for such legislation as will give most effective embodiment to this idealism.

FR. DE HOVRE. *German and English Education.* New York: Char es Scribner's Sons, 1917. Pp. 108.

This is a stimulating comparative study of the two civilizations now at death grips with each other, and of the type of education upon which each is based. The author, who was professor of the philosophy of education at the University of Louvain, has had exceptional opportunities for studying both systems at close range. He is clearly cognizant of all the magnificant achievements of German education in moulding the thoughts and will of the people, yet he claims that in point of organic principles England has nothing to learn from Germany. Reforms are demanded, and the effect of these reforms may be far-reaching, but they must be worked out on the basis of fundamental English characteristics and ideals.

E. A. DOLL. *Clinical Studies in Feeble-Mindedness.* Boston: Richard G. Badger, 1917. Pp. 232. $2.50.

This is an account of a series of studies selected from the great mass of material accumulated at Vineland to illustrate and clarify some of the fundamental questions about feeble-mindedness. The author has aimed to indicate the several points of view from which feeble-mindedness may be defined and to formulate such definitions, to illustrate differential diagnosis of mental defect, and to evaluate typical examinational data. According to the psychological criterion feeble-mindedness is arrested development of general intelligence at a point not exceeding the capacity of a twelve-year-old normal child. The social criterion calls for social inefficiency due to arrested development. The pedagogical criterion rests upon scholastic deficiency. The medical criterion assumes a diseased or defective condition characterized by a greater or less number of physical abnormalities, particularly of the nervous system. The somatic criterion recognizes a more or less complete retardation of the entire organism, physical, physiological and anatomical, as well as mental. The hereditary criterion asserts that true feeble-mindedness results from deficiencies in the germ plasm, and that these deficiencies will show themselves in the descendants in definite relations. The application of these criteria is illustrated by the detailed study of six typical cases. The book will be carefully studied by those interested in the diagnosis and treatment of feeble-mindedness.

JAMES DREVER. *Instinct in Man. A Contribution to the Psychology of Education.* New York: G. P. Putnam's Sons, 1917. Pp. x, 281. $2.25.

This is an elaborate and systematic study of a subject that has played a very important role in works on educational psychology in the past two decades. As such it deserves thoughtful and critical consideration. The author first presents a historical survey of philosophical and scientific views of instinct from Hobbes to the present day, then enters upon a psychological analysis of the nature of instinct in its relation to intelligence and meaning, and finally traces the relation of instinct to other aspects of experience, notably the emotions, and to the phenomena of de-

velopment and education. He seems disposed to reject Thorndike's invitation "to discard such vague phrases as the instincts of pugnacity, gregariousness, cruelty, curiosity, constructiveness, play and the like, and to substitute for them descriptions of the definite responses which are made to definite situations." His reasons for this rejection are that in the first place the definiteness of the responses is illusory, and in the second place that "with a view to the development of a psychology of education the 'class names' are exceedingly valuable, since their very 'vagueness' indicates that indeterminateness which is for the educator so significant a feature of the instinctive equipment of the human being." After this it is not surprising that McDougall's views of the nature of instinct are accepted, and that the discussion is devoted largely to the verbal and schematic analyses of James, Morgan, Shand, Groos and others.

S. FERENCZI. *Contributions to Psycho-Analysis.* Authorized translation by Dr. Ernest Jones. Boston: Richard G. Badger, 1916. Pp. 288. $3.00.

The author is medical advisor to the Hungarian law courts, and has made extensive studies in the field of psycho-analysis. In this book the translator presents fifteen of his more important papers. They offer brilliant examples of the interpretation of psychopathic disorders in terms of sexual sublimations, such as we have learned to expect from the psycho-analytic school.

HENRY T. FINCK. *Richard Struass, The Man and His Works.* Boston: Little, Brown and Company, 1917. Pp. xxvii, 328. $2.5C.

Music lovers will be keenly interested in this biographical and critical study of the greatest living master of orchestral tone coloring made by the veteran musical critic and biographer of Richard Wagner. In view of the author's admiration for Wagner it is perhaps not surprising that he should take a somewhat derogatory attitude toward Strauss's opera technique. The book is indeed no outpouring of fulsome adulation. So coolly critical is it that many of Strauss's admirers will resent it as another instance of the failure of the old to understand and appreciate the new. Yet in spite of its many sharp critical thrusts, and its perhaps ungenerous comparisons with "the great Richard," the book sets forth Strauss the musician in clear, vigorous, forceful outlines, and makes us realize that here we meet a musical genius of the first order. No thoughtful student of music as a factor in life can afford to overlook this book.

THOMAS WALTON GALLOWAY. *The Use of Motives in Teaching Morals and Religion.* Boston: The Pilgrim Press, 1917. Pp. xi, 187. $1.25.

This little manual is an attempt by a successful teacher and textbook writer of zoology to apply to religious teaching the principles that have been derived in the teaching of school subjects. In the introductory chapters he outlines and illustrates these principles, and in chapters four to nine he applies them to the motivation of religious and moral teaching. The concluding chapters deal with the forms of expression in which such teaching should eventuate. The book is well conceived and skillfully executed, and should be of great assistance in training the immature teachers of the Sunday School.

FR. AGOSTINO GEMELLI. *Il nostro Soldato. Saggi di Psicologia Militare.* Milano: Societa Editoriale Vita e Pensiero, 1917. Pp. xii, 339.

In Europe the immediacy and long continuation of the war have produced a more profound and wide-spread effect upon all phases of thought than has yet been the case in this country. European medical, educational and psychological discussions are all dominated by it. The present book is an interesting and serious study of the psychology of the soldier. Among the topics considered are the psychic factors entering into victory, the mental attitude of trench life, the intellectual and moral aspects of battle, a psychological analysis of courage on the battle-field, fear, superstition, military folk-lore, the songs of the soldiers, the acoustic illusions of battle, the psychology of the officer, psychology and strategy, and the pathological effects of fear in war. The author is a captain in the medical service, and has had an excellent opportunity for first hand studies of the conditions which he discusses. The

book abounds in shrewd observations and reflective explanations. It is not merely a record of personal experiences, but a well-digested scientific work in which the author takes due account of recent French and English contributions to the topic of which he treats.

FRED. M. GERLACH. *Vocabulary Studies.* Colorado College Studies in Education and Psychology, Number one, 1917. Pp. 123. Thirty-five cents.

In view of the fact that language is at once the highest manifestation of human activity and the most important means of conserving and advancing human culture through education, it is surprising that so little attention has been devoted to the scientific study of its acquisition. The author's general discussion summarizes the scattered studies that have been made, and presents in comparative tables the results obtained by Kirkpatrick, Babbitt, Doran, and others. It is remarkable that no mention is made of Terman's excellent studies in this field. The third part gives a detailed account of the "false definition test" as employed at Colorado College. One thousand words were selected "by position" from Funk and Wagnalls' New Standard Dictionary representing a total of 250,000 vocabulary terms. Four hundred of the selected words were so unfamiliar that the subjects were merely asked to check those with which they were acquainted. Each of the remaining six hundred words was supplied with four definitions, only one of which was correct, and the subjects were asked to check the correct definitions. Both of these lists are given in appendices, together with a representative list of 100 words arranged in order of difficulty. The lists were given to students ranging from high school freshmen to college seniors, and the average computed vocabulary varied from 66,650 for the high school freshmen to 94,075 for the college seniors. These results are over four times as large as those secured by Kirkpatrick. There is a bibliography of 63 numbers.

CLARENCE TRUMAN GRAY. *Types of Reading Ability as Exhibited through Tests and Laboratory Experiments.* Supplementary Educational Monographs, No. 5. Chicago: University of Chicago Press, 1917. Pp. xviii, 196. $1.25.

This is a detailed study of the motor phases of reading, of the span of visual perception, and of the measurement of abilities in oral and silent reading on the same subjects. With the most refined and delicate apparatus the eye-movements in reading were photographed, the subject's breathing was recorded, the vocal utterances were fixed on a phonograph synchronized with the eye-movement apparatus, and the range of perception was determined by short exposure, recognition and Aussage tests. Good and poor readers were studied from each of the elementary grades, and from high school and college classes, and their characteristic differences noted. In the final section there is a study of the effect of practice upon certain difficulties in reading. The practice included speed training, increase in span of attention, training in comprehension, reduction of vocalization, and training in phonics.

ALBERT BUSHNELL HART. *New American History.* New York: American Book Company, 1917. Pp. vii, 650, lii.

This is a splendid example of the new type of history that is now being offered to the boys and girls in the high school. The product of the ripest and most authoritative scholarship, the book is written in a simple narrative style, richly illustrated with attractive pictures, drawings and maps. and amply furnished with references and helps for study. Of the entire 650 pages 450 are devoted to the development of the country since the close of the Revolution. Particular stress is laid upon the social forces that were active in that development. The general use of such a text will go far to develop that broad spirit of Americanism of which we now stand so much in need.

CHEESMAN A. HERRICK. *History of Commerce and Industry.* New York: The Macmillan Company, 1917. Pp. xxv, 562. $1.60.

The author's aim in this book has been to present the essentials of history from the commercial and industrial point of view, to trace the history of a great movement

rather than the history of nations. Beginning with the Nile and the Euphrates the currents of trade are followed in the development of Phoenecia, Greece, Rome, the Italian republics, the Saracens and the Crusaders, the markets and fairs of the middle ages, the Hanseatic towns, the spirit of exploration in the fifteenth and sixteenth centuries, colonial Spain, commercial England, Dutch commerce, English North America, the industrial revolution, modern Britain, German commercial and industrial expansion, Russia, China, Japan, Latin America, the United States as a world power, and the outlook for a greater America. This is a type of history that is vastly more significant than military campaigns or the lives of individual rulers. The story is well told, and illustrated with many striking photographs. There are excellent bibliographies and suggestions for study.

James L. Hughes. *Training the Children.* New York: The A. S. Barnes Company, 1917. Pp. 148.

The purpose of the author is to contrast the old training of coercion with the new training of freedom and self-activity. The book is the outgrowth of a long life of teaching and supervision, and yet shows nothing of the cut-and-dried attitude of the administrator, but is full of sympathy with and understanding of the child. "Freedom for the child should mean freedom from abuse, freedom from coercion, freedom to work out his own plans in his own life department."

A. Eaglefield Hull. *Scriabin, A Great Russian Tone-Poet.* New York: E. P. Dutton and Company, 1916. Pp. viii, 304. $1.25.

Scriabin is a composer of whom all too little is known in America. Of an original, aggressive disposition, he found it difficult to work in harmony with others, and thus failed to receive in his time the recognition that was his due. Born in 1871, he was trained as an officer in the Russian army, but early manifested a passion for music and received instruction under some of the most celebrated Russian composers. He became a great piano virtuoso, and was an extremely prolific composer for that instrument and for the orchestra. He might well be called the Russian Chopin. His music is characterized by wonderful richness of harmony and a profound mysticism which developed into an attempt at the harmonization of all the arts. The author of the present book distinctly disclaims the intention of writing a biography, but he succeeds admirably in giving us a clear-cut impression of the genius of the artist and of the rich store of music which he produced before his untimely death.

Louis Huot et Paul Voivenel. *Le Courage.* Paris: Librairie Felix Alcan, 1917. Pp. vii, 358. Francs 3.50.

This is another of the excellent books on the psychology of the war which are now appearing in Europe. According to the authors "war has been and will be," and it is the duty of the scientific student to examine its fundamental nature and analyze the phenomena which it reveals. Both the authors are medical officers in the army and have seen much active service. They claim that courage does not exist among animals, that real courage requires a development of intelligence that is found only in man. They trace the natural history of courage, describe the fear that is keenly recognized but held in check by steadfastness of purpose, the force of altruistic ideas, the effects of habit and experience, the love of risk, and depict the characteristics of group courage, of national spirit, of the intoxication of battle, and of the pathology of courage.

Alexander Inglis. *Principles of Secondary Education.* Boston: Houghton Mifflin Company, 1918. Pp. xvi, 741.

This book is a veritable encyclopedia of secondary education, but differs from the treatises of Johnston and Monroe in the unity of aim and of construction secured by a single author. Part I deals with the pupils, and surveys the physical and mental traits of high school pupils, their individual differences, and the general character and classification of the high school population. Part II discusses the high school as an institution, sketches the history of secondary education in America and other countries, shows its relation to elementary and higher education, and develops its social principles and fundamental aims. Part III describes the means and materials

of secondary education, the program of studies, and the place of the special subjects, as English, foreign languages, mathematics, natural sciences, social sciences, practical and fine arts, etc., in that program. The treatment of each of these topics is remarkably well-rounded and complete.

HERMAN G. JAMES. *Municipal Functions.* New York: D. Appleton and Company, 1917. Pp. xi, 369. $2.00.

The purpose of this book is to inform the average citizen in regard to what is being done by progressive municipalities, and to point the way toward a fuller utilization of municipal activities for the welfare of all. Among the topics considered are public safety, health, education, morals, social welfare, city planning, public works, municipal ownership, and municipal finances. If educators are to fill the places of community leaders, they need to give careful consideration to municipal problems, and this book will be of great assistance to them.

BUFORD JENNETTE JOHNSON. *Experimental Study of Motor Abilities of Children in the Primary Grades.* Johns Hopkins University Studies in Education, No. 2. Baltimore: The Johns Hopkins Press, 1917. Pp. 62. Seventy-five cents.

This study of the motor activities of children was suggested by the success with which motion studies have been utilized in increasing the efficiency of industrial workers. The author gives a brief summary of previous studies on the speed and accuracy of movement, and recounts experiments with nineteen first grade children in rate of tapping, steadiness of control, reaction times, accuracy in throwing, and preference for side of the body used. There is marked improvement with practice, but the improvement is so irregular that it is not safe to take the first trial or the average of the first few trials as the index of ability.

EDWARD SAFFORD JONES. *Influence of Age and Experience on Correlations Concerned with Mental Testing.* Baltimore: Warwick and York, 1917. Educational Psychology Monographs, No. 22. Pp. 89. $1.25.

This study is the outcome of work done in connection with the Cincinnati Vocational Bureau. Two hundred and three unselected boys were examined when they secured working papers at the age of fourteen and each year thereafter for four successive years. From a large number of mental and physical tests the following were chosen for this study: Cancellation, substitution, rote memory, completion and opposites tests. The author finds a marked fidelity to intellectual type throughout the adolescent period, and believes that one well-rounded testing of an individual is likely to place his general intellectual rank for several years to come.

TRUMAN LEE KELLEY. *Mental Aspects of Delinquency.* University of Texas Bulletin, March 1, 1917. Pp. 125.

Two hundred and ninety-three delinquent boys in the Texas State Juvenile Training School were examined for their height, weight, vital capacity, head measurements, pubertal development, vision, hearing, nose and throat, strength of grip, handedness, tapping, school grade, Binet age, constructive ability, and ability in a modified Trabue completion test. The boys were found to be normal or above in height, weight, vital capacity, and head circumference, but below normal in all the other measurements. They were conspicuously low in the mental tests. There are detailed tables of the individual results, a discussion of the findings on the basis of other studies of delinquency, and a valuable annotated bibliography.

MABEL HYDE KITTREDGE. *The Home and its Management.* New York: The Century Company, 1917. Pp. 385. $1.50.

This "Handbook in Home-making" with its three hundred inexpensive receipts is very timely in view of the obligation resting upon us to economize both food and money to win the war. The book is designed not for the university graduate, but for the plain girl in the kitchen. The suggestions for the selection and preparation of foods will be valuable even to those who have had wide experience in housekeeping.

M. G. Kirkpatrick. *The Rural School from Within.* Philadelphia: J. B. Lippincott Company, 1917. Pp. 303.

Personal experiences are always interesting, and this account of personal experiences in country school teaching is no exception to the rule. In some respects no type of teaching is more difficult, makes greater demands upon the patience, steadfastness and ingenuity of the teacher. The young and inexperienced teacher will find many helpful hints in these pages.

Daniel Wolford La Rue. *The Science and the Art of Teaching.* New York: The American Book Company, 1917. Pp. 336.

Young teachers need to be encouraged to survey their work as a whole in order to see the individual tasks in their proper perspective. Methods of teaching are determined by four factors: The nature of the child (involving a study of general and individual psychology), the manipulations of the teacher, the nature of the subject matter taught, and the educational ideals of the time and the community. In the discussion of the second factor the author includes a brief chapter on pedagogical measurements.

Darwin Oliver Lyon. *Memory and the Learning Process.* Baltimore: Warwick and York, 1917. Pp. 184. $2.00.

This elaborate investigation of memory in connection with education gives an account of experiments which the author has been conducting since 1906. At the outbreak of the war, in 1914, the manuscript was lost in Germany, and the book had to be rewritten from the author's original notes. The discussion includes an analysis of the mental activities included under "memory," the subdivisions of the topic, with especial reference to the learning process, the relation of length of material to time of learning, the relation of quickness of learning to retentiveness, and the educational value of psychological research, with a critical consideration of the value of mnemonic systems. Appendices present the details of the tests, and a pocket at the back contains a number of finely drawn plates in colors. It is a valuable contribution to the study of memory, and will command wide attention.

Cyrus D. Mead. *An Experiment in the Fundamentals.* School Efficiency Monographs. Yonkers-on-Hudson: World Book Company, 1917. Pp. xiv, 54. Sixty cents.

A comparative study of the Courtis Standard Practice Tests and the Thompson Minimum Essentials on nine hundred fifth grade pupils, divided into approximately equal groups. The results of the study were in favor of the Courtis Practice Tests, the classes using these showing almost twice as much gross gain as those using the Thompson material. The reason for this superiority seemed to lie in the greater zeal aroused in the pupils, and the greater ease in advancing and controlling the pupils' progress.

Edwin L. Miller. *English Literature: An Introduction and Guide to the Best English Books.* Philadelphia: J. B. Lippincott Company, 1917. Pp. 597.

Most high school courses in English literature demand the worst type of memory grind. Page after page of critical comment on the works of authors the pupils have never read and have no interest in must be prepared and recited. The present text is constructed on a different plan. It is intended as a guide to reading, and the effort of the author is to stimulate a desire to become familiar with these literary masterpieces that make an appeal to every properly constituted youth. Thus he takes the pupil's point of view, and presents his case so appealingly that one is consumed with eagerness to read these good things. Let us hope that teachers will use the book as the author intends, and will not force it into a recitation manual.

Walter Scott Monroe, and Others. *Educational Tests and Measurements.* Boston: Houghton Mifflin Company, 1917. Pp. xxii, 309. $1.50.

A real science of education would enable us to determine just where each pupil stands at the beginning of the term in each school subject, and what progress may be expected of him during the term as a result of a given procedure. We are far

from this goal as yet, but the progress made toward it in the past five years is well set forth in the present volume. We find here an excellent survey of the Courtis, Stone, Woody, Cleveland, and other tests in arithmetic, the Thorndike, Haggerty, Starch Courtis, Brown, Gray, and Kansas reading tests, the Ayres, Buckingham and Starch spelling scales, the Thorndike, Ayres, Johnston, Breed, Freeman, and other handwriting scales, the Hillegas, Harvard-Newton, Breed, Willing, Trabue, Starch, and other language scales, and the fragmentary work that has been done in the study of the various high school subjects. There is a chapter on statistical methods, and one on the use of standard tests in the supervision of instruction.

DeWitt H. Parker. *The Self and Nature.* Cambridge, Mass.: Harvard University Press, 1917. Pp. ix, 316. $2.00.

However emphatically we may be told that this is a materialistic and grossly practical age, we still find a goodly number of those enterprising souls to whom the romance of metaphysics makes a powerful appeal. It is in the spirit of a quest, of an intellectual adventure, of a logical voyage of discovery that the author bids us set out with him in an examination of the basic facts of experience. On our quest we meet and tilt with many familiar characters, James, Royce, Bergson, Bradley, Russell and others, and these passages at arms are attractive features of the adventure. Do we arrive anywhere, do we discover anything? This is an individual question, and the answer must be left to the individual reader. At least it is a very entertaining and stimulating journey.

Angelo Patri. *A Schoolmaster of the Great City.* New York: The Macmillan Company, 1917. Pp. 221. $1.25.

This is one of the most fascinating and inspiring books on general educational topics that the past year has furnished. The son of a poor Italian immigrant, starting to school only when eleven years old, the author has pushed forward until he has become one of the most successful and influential school principals in New York City. It is not a consecutive autobiography that he gives, but rather a series of dramatic pictures, portraying the strivings and aspirations of a sincere and earnest teacher in his efforts to touch and quicken the lives of the children in a great city. It depicts for modern New York conditions something of the same spirit and courage and kindliness that is shown for the middle west three-quarters of a century ago in the life-story of Abraham Lincoln.

Rudolf Pintner and Margaret M. Anderson. *The Picture Completion Test.* Educational Psychology Monographs, No. 20. Baltimore: Warwick and York, 1917. Pp. 101. $1.25.

This is a study of the application of a single test, the Healy Picture Completion Test, to over 1500 children from five to sixteen years of age. The test is described, previous uses of it discussed, the results are tabulated, the method of scoring worked out, norms for each age are derived, and there is a consideration of the relation between the results and sex, social status and school standing. The accelerated pupils did better than the retarded pupils, showing that the test differentiates between good and poor pupils. Such detailed studies as this contribute materially to the advancement of the science of mental testing.

W. F. Robie. *Rational Sex Ethics.* Boston: Richard G. Badger, 1917. Pp. 356. $3.50.

In every civilized country men are now seriously concerned with the birth-rate. But the birth-rate depends upon economic, social and psychological factors which underlie the founding of a family. The present book is the sanest and frankest consideration of one of those factors, sexual attractiveness, that we have ever seen. On the basis of a large number of case studies the author, who is a practicing physician, gives valuable advice for the maintainance of a normal, healthy sex life. While the book is for sale only to members of the medical and legal professions, no physician could do better service than present a copy to every young married couple of his acquaintance.

WILLIAM M. SALTER. *Nietzsche, the Thinker.* New York: Henry Holt and Company, 1917. Pp. x, 539. $3.50.

At the present time Nietzsche is more or less justly held responsible for the frightfulness and ruthlessness with which Germany is waging war. Yet while the Germans may find philosophical justification for much of their conduct in Nietzsche's writings, it must be remembered that Prussianism and the whole military system for which Prussia stands were utterly abhorred by the philosopher. The present admirable and thoroughly scholarly work gives a clear and coherent account of Nietzsche's philosophy and an especially valuable critique of his theory of morality. The fragmentary, aphoristic style and lack of systematic development in Nietzsche's writings make such an organizing commentary necessary for an understanding of his fundamental doctrines. The author has acquitted himself masterfully, and his book will stand as the authorative interpretation of Nietzsche in English.

OSIAS L. SCHWARZ. *General Types of Superior Men.* Boston: Richard G. Badger, 1917. Pp. 435. $2.50.

In a preface Jack London says of this book: "General Types of Superior Men is one of those immortal, epoch-making works which appear only at very long intervals . . . . . . The psychology of the various and many types of superior men is merely the nucleus of his subject from which Mr. Schwarz ventures into all regions of human knowledge in order to build up his original philosophy of human life." A less sympathetic commentator might characterize the book as a Bolshevik attack on what is ordinarily called the "superior man." Of psychology in any scientific sense the book shows not a trace.

Science and Learning in France, with a Survey of Opportunity for American Students in French Universities. An Appreciation by American Scholars. Published by the Society for American Fellowships in France, 1917. Pp. xxxviii, 454.

This sumptuous volume was prepared to show what France has done for the advancement of science, and what American students might gain by study in France. Each chapter sets forth briefly what noteworthy contributions Frenchmen have made to a given field of human knowledge, who are the leading French contributors to that field at the present time, and what opportunities are offered for graduate work in that subject in France. The book gives a good account of past and contemporary French science.

W. T. SEDGWICK AND H. W. TYLER. *A Short History of Science.* New York: The Macmillan Company, 1917. Pp. xv, 474. $2.50.

This book is the outgrowth of a lecture course which the authors have given for several years to students of the Massachusetts Institute of Technology. It is distinctly more elaborate and contains more abundant quotations from source material than other recent texts on the history of science. Of the scientific accuracy of a book by such authors there could be no question. One misses, however, the imaginative touch which fills in the gaps, rounds out the probable trend of development, and gives the student the interpretive benefit of the author's study. Such a procedure is not scientifically accurate, but it inspires the student and helps him recreate the past. The book is richly illustrated with portraits of eminent scientists and photographs of maps, diagrams, and instruments.

ELIZABETH SEVERN. *The Psychology of Behavior.* New York: Dodd, Mead, and Company, 1917. Pp. ix, 349. $1.50.

The sub-title of this book is "A Practical Study of Human Personality and Conduct with Special Reference to Methods of Development." The topics treated are the subconscious, intellect (perception), imagination and memory, will, emotion, sex and the ego. In spite of the somewhat formal analytical tendency in the discussion of these topics, the author's positions are well taken, and there are many shrewd practical observations that are commendable.

FRANK CHAPMAN SHARP. *Education for Character. Moral Training in the School and Home.* Indianapolis: The Bobbs-Merrill Company, 1917. Pp. 453. $1.25.

The central problem of this book is how directly to develop and strengthen loyalty

to moral ideals as such. The author confines himself to a survey of the concrete aims of moral education and the agencies and instruments by which they may be attained. There is a good discussion of the importance and limitations of the personality of the teacher, and the factor of personality in leadership among the pupils. The advantages and disadvantages of pupil government are pointed out, the need for direct moral instruction is emphasized, and a somewhat elaborate plan for moral instruction on the basis of biography and social service is worked out. Appendices contain a detailed program of moral instruction for the elementary school, a series of practical questions and exercises, and an extensive bibliography.

FRANK WEBSTER SMITH. *The High School. A Study of Origins and Tendencies.* New York: Sturgis and Walton Company, 1917. Pp. xviii, 458. $2.00.

The high school by its wonderfully rapid development and its articulation with both the elementary school and the college is one of the most interesting and significant aspects of American education. Impressed with this fact the author set himself to inquire into the history of the institution, and the result is a valuable historical study of secondary education from primitive times to the present. There is an excellent account of Quintilian's secondary school, secondary education in medieval times, a briefer sketch of secondary education from the fifteenth to the nineteenth centuries, and a trenchant discussion of high school practices and tendencies at the present time. The book is a valuable addition to the literature of secondary education.

LYNN THORNDIKE. *The History of Medieval Europe.* Boston: Houghton Mifflin Company, 1917. Pp. xxi, 682.

It has become the custom in many colleges to make the first course in history a history of medieval Europe. A study of this period seems to lay the best foundation for subsequent courses both in modern and in ancient history. It is to provide a text for such a course that the present book is intended, and casual inspection leads one to think that it will serve its purpose excellently. The discussion is much fuller and more detailed than the ordinary text, and much attention is paid to social and economic conditions, as the City of God, the spread of monasticism, the feudal land system, the medieval church, the rise of towns and guilds, medieval literature, and the Italian renaissance.

J. E. WALLACE WALLIN. *Problems of Subnormality.* Yonkers-on-Hudson: World Book Company, 1917. Pp. xv, 485. $3.00.

Undoubtedly the subnormal child has had more than his just share of attention in the educational discussions of the past four or five years. But with all the consideration given to him we are still far from giving him the kind of treatment that he needs, and we are even farther from relieving the normal child from the drag of his company. In the first hundred pages of this book the author sketches the history of the movement, and reports a study of 170 pupils in special classes in St. Louis. In chapter two he considers the problem of diagnosis, reviews the reports of feeble-mindedness resulting from the surveys of several institutions, and cautions against a pronouncement of feeble-mindedness on the basis of any single test, such as that of Binet. Chapter three lays down thirteen principles for the organization of work with mentally retarded children, and the remaining four chapters treat of feeblemindedness in its educational and social aspects, epilepsy, state provisions for defective children, and the hygiene of eugenic generation.

MRS. HUMPHREY WARD. *Missing.* New York: Dodd, Mead and Company, 1917. Pp. 398. $1.50.

This is an English romance of the early days of the war. The hero, an officer who has been slightly wounded, and is now on leave, meets, falls in love with, and marries the heroine in a few weeks. After a short honeymoon he is recalled to the colors, and the young wife is left to the machinations of a rather calculating sister who does her best to establish friendly relations with a neighboring squire. Later the husband is reported "missing" and in the absence of further news the squire presses his suit, and almost succeeds in comforting the 'widow,' when news comes that the

husband has been found severely wounded. In deep remorse the wife hurries to her husband, and although he subsequently dies, she will have nothing more to do with her sister or the squire. There are many skillful characterizations of the cross currents of English thought at the outbreak of the war.

HUTTON WEBSTER. *Early European History.* New York: D. C. Heath and Company, 1917. Pp. xxxv, 715. $1.60.

In less than 300 pages the author reveals to high school pupils the vital experiences of the ancient peoples. The greater part of the space is devoted to the period from the fifth to the seventeenth century. In treating this period prominence is given to the rich social lessons, the slow emerging of the common people, the progress of the ideals that sway men and nations. The volume was planned in accordance with the recommendations made in 1916 by the Committee on Social Studies of the Commission on the Reorganization of Secondary Education of the National Education Association. It is written in a clear, attractive style, is well equipped with maps and illustrations, and is well suited to the interests and capacities of high school pupils.

H. B. WILSON. *Training Pupils to Study.* Baltimore: Warwick and York, 1917. Pp. viii, 72. Fifty cents.

This little book is the outgrowth of the efforts of the teachers of a school system to teach their pupils how to study. The author analyzes study into eleven essential factors, abundantly illustrating each point with instances from schoolroom practice. This is followed by reports on training pupils to study in the various school subjects from the second to the eighth grade. While these reports only indicate a few of the things that might be done, they at least show what some enterprising teachers are doing to stimulate study.

VOLUME IX APRIL, 1918 NUMBER 4

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

AN EXPERIMENTAL STUDY OF METHODS IN TEACHING HIGH SCHOOL CHEMISTRY

WILLIAM H. WILEY
Miami University, Oxford, Ohio

No field of secondary education presents more numerous or more important problems than the teaching of the sciences. The age is a scientific one. The applications of science are advancing by leaps and bounds. Yet the teaching of science in the high school is not having the influence on our boys and girls that it should. In the field of chemistry in particular no experimental studies have been made either on subject matter or on methods of teaching. In a recent article Professor E. B. Spear[1] points out many of the unanswered problems that face the chemistry teacher in teaching the simplest things. He inquires what would be the correct amount of time to spend on chemistry and how should this time be distributed. Is a pupil justified in taking one year of chemistry and then stopping? In the process of learning chemistry, involving, as it does, manipulation, observation, correlation of ideas, logical memory, imagination and reasoning, there are many problems from which Spear suggests the following. How can we best promote dexterity in the manipulation of apparatus? Does observation gained in chemistry increase accuracy of observation in the ordinary circumstances of life? Should careful directions be given to the student for observation, or should he be free to observe what he will? Should pupils correlate facts with facts, theories with theories or facts and theory? Should the instruction be largely a correlation of facts, theories and principles? Should the experiments be chosen of simply to illustrate a point or to correlate the new facts with those previously learned? What is the best way to teach relationships? In logical reasoning should we insist upon the student holding to

[1]E. B. SPEAR. *Problems in the Experimental Pedagogy of Chemistry.* Journal of Educational Psychology, Vol. 6: 1915, 231–241.

conclusions limited to the data at hand or generalizing beyond these data? How are we to teach the pupil to test his premises? Should the student know before-hand the purpose of his experiment or should the results determine the purpose? Is not the pupil often prejudiced by the purpose? Should chemistry present constructive problems to train the imagination? Would it be of any value to have the pupils predict what they would expect from an experiment? Is it better to have pupils construct apparatus from meager descriptions or to give them full directions and sketches? Spear thinks that these problems belong to the field of experimental pedagogy, and that their solution would greatly enhance the effectiveness of chemistry teaching.

There are many other problems in the teaching of chemistry. We ought to know what material is the best for beginning pupils, the kind and amount that can be taught effectively in the high school course of one year. We ought to know whether the training obtained in a high school course justifies itself, or whether we ought to spend the time teaching valuable facts. No matter how we look at the question of high school chemistry there is one point that we must consider. The learning of facts is one of the important functions of chemistry teaching. Whether the pupil is to form habits, to observe, or to reason, memory plays a most significant part. The learning of fundamental facts and principles seems to be the first problem for consideration.

This investigation will, therefore, attempt to apply the methods of experimental pedagogy to the problem of teaching certain facts in chemistry. *The main problem* of this study is to determine the best of three methods of teaching chemistry, the text-book recitation method, the so-called lecture method, and the laboratory method. By the best method we mean that which gives the best results in the least possible time and with the least expenditure of energy. The results are then to be determined qualitatively and quantitatively as well as their effect upon the student.

By the *text-book method* we mean the study of the book on the part of the pupil without apparatus or other helps. After the pupil has studied the lesson he recited it to the teacher according to the latter's questioning.

By the *lecture method* is meant the teaching of the lesson by alternately talking and questioning the pupils concerning a topic they have not previously prepared. In this method apparatus and

other helps are used and the pupils' attention is centered upon the demonstration performed by the teacher. This is commonly called the informal lecture.

By the *laboratory method* the pupils are given a problem and directions as to its solution. They are given the necessary apparatus and materials and allowed to work. While they are working, the teacher continually questions them to keep the problem before them. They may write notes, but this is not required for these experiments. Learning here means both immediate and permanent retention of the facts imparted in the lesson. As far as these tests go we must limit our measure of retention to one month after the test or lesson is taught.

Secondary Problems: (1) The investigation of permanent retention as well as immediate recall. (2) The study of sex differences in learning chemistry. (3) The study of individual differences in learning chemistry. (4) The effect of the different methods on the pupil, in regard to interest, attention and fatigue. The criterion for judging these problems will be the pupils' efficiency in responding to written tests.

THE TESTS

For the preliminary tests a number of lessons were used, some selected from text books and others written by the author. The most important of these was "The Chemistry of the Parlor Match." The final tests consisted of three lessons written by the author. The problems for the lessons were as follows:

Lesson I. What makes things burn?
Lesson II. Why is baking powder used in cake?
Lesson III. How is cloth colored?

Each lesson was so written as to contain the important ideas about each problem named. The lessons were of practically the same length and of the same difficulty in so far as this could be determined by a preliminary test. Each lesson was typewritten on uniform paper 8½″ x 11″ and each student had a copy. As far as possible the papers were uniform in size, difficulty and value of ideas.

In the laboratory method each pupil was supplied with all the necessary materials and apparatus to work out the problem. In the lecture method I used apparatus in teaching each of the lessons and drew the facts from their observation. Additional facts I told them, until they had all the ideas in the lesson.

The Subjects

The subjects for these experiments were 24 pupils in the McGuffy High School of Miami University. They were in the junior and senior years of high school and all had studied physics. They were probably a little above the average high school pupil in intelligence owing to their environment and training. They ranged in age from 16 to 20. None of the pupils had studied chemistry and none of the work was familiar to them. In grouping the students for the tests it was necessary to use their physics marks as a partial criterion. These with a set of preliminary marks were averaged and on this basis the 24 pupils were divided into three groups of eight each. Each group was assumed to have the same average ability in so far as marks form a measure of ability. The pupils were taught in groups of eight each and tested immediately after the lesson. The groups were again tested one week after the lesson and again four weeks after the lesson. The tests were in writing and were given under uniform conditions as far as these could be controlled.

Preliminary Tests

A number of preliminary tests were given in order to determine the advisability of using certain methods of procedure in these tests. Thus a lesson was selected from a book on chemistry and the pupils told to study it until they knew it. They were asked to record the time required to learn it. Other tests were written out in the form of laboratory directions and pupils were asked to carry out these in the laboratory. They were then asked a number of questions to which they were to write out the answers. The unsatisfactory results of these experiments led the writer to control the entire lesson as well as the tests, for it was found impractical to depend upon the student to carry out the directions.

A second set of preliminary examinations was given to determine the relative value of the lessons to be taught and secondly to determine the average ability of the class. For the first test five students were selected who had no knowledge of chemistry. Copies of each lesson were given the group and they were told to read the lesson carefully once. They were then given paper and told to write all they could remember of the lesson. On the succeeding day at the same time they followed the same procedure with lesson two. On the next day they did lesson three. The papers were collected and carefully graded.

Now in order to determine roughly the ability of the members of the class another lesson (The Chemistry of the Parlor Match) was presented to the class. Each member was given a sheet of the printed lesson. I carefully read the sheet to them, and then told them to read it carefully once and as soon as they finished to turn over the paper. They were then asked to write out on a sheet of paper all they could remember of the lesson. These papers were collected and carefully graded.

Final Tests

1. Plan of the Tests.

Each of the three lessons were typewritten on paper 8½″ x 11″ and each pupil was provided with a copy of each lesson. The lessons and tests were all given at the same hour, at 10:45 A. M. on Tuesday, Thursday and Friday of each week. In so far as the conditions of environment and teaching were controllable they were made uniform.

The Text-Book Recitation Lesson

In this lesson the pupils were each given a copy of the lesson and were told to study it carefully until they felt they knew it. Anything not understood was explained. When the majority stated that they knew it, all copies were laid down and I asked various questions concerning the lesson. In this way I covered the important points in the lesson. I also answered any questions they asked about it. The time was recorded. I now passed paper and asked them to write out a complete resume of the lesson. These papers were collected and graded. One week later the same class were given paper and asked to write out the lesson they had had last week. Nothing was said at the first test that there would be a further test. Likewise nothing was said of the third test, which was given like the others at the end of four weeks. All papers were graded according to the method described below.

The Lecture Method

In this lesson the materials and apparatus were arranged on the table before me. I presented a problem to the class, *e. g.* "What makes things burn?" and informally discussed the material on one of the typewritten sheets. I talked most of the time but occasionally asked a question which I thought they could answer. While talking and telling them about burning, oxygen and so on, I performed the experiment to illustrate the subject matter on the sheet

of lesson II. I wrote new and difficult words on the blackboard and at the end summarized the important points in the lesson. I then gave them paper and asked them to write out all they remembered of what I said. I collected and graded the test papers. A second and a third test were given one week and four weeks later, just as described in the lesson above.

The Laboratory Method

In this method the pupils were taken into a laboratory and each assigned a definite place. They were then given a problem and a sheet containing the directions for the lesson. They were told to read the lesson and in the meantime I placed the necessary material on the desk. As I passed the materials I asked them what they needed and as they told me I gave it to them. They then went to work and performed the experiment. I stood by and watched each student carefully to see if he knew what he was about. I continually asked questions of different members of the class to be sure that the main points of the lesson were being established. When they concluded the experimenting I had some of them summarize the points of the lesson and show me how they had solved their problem. I then passed paper and told them to write out the lesson as they remembered it. A week and a month later I had them write out the lesson again without any preparation in the meantime. They had no idea of the second and third tests and this removed any incentive to look up the lesson in the meantime. All the tests were collected and graded as follows.

2. Method of Scoring.

Each of the three lessons was carefully analyzed by Professor Whitcomb (Head of the Chemistry department, Miami University) and myself, and the lessons were divided into a definite number of important ideas and a second group of unimportant or amplifying ideas. The former were given two credits each and the latter were given one credit each. The lessons were then revised so that the value of all the credits in each lesson should total 40. The distribution of these credits for each lesson was as follows:

	Lesson I.	Lesson II.	Lesson III.
2 credits	13—26	14—28	13—26
1 credit	14—14	12—12	14—14
Total	40	40	40

In correcting the papers full value was given for the idea correctly stated, half credit for partly correct and no credit if missing. If an incorrect idea was stated for the correct idea one credit was deducted from the total. Ideas not in the original lesson were not credited.

Results

The following tables give the numerical results of the preliminary tests.

a. Preliminary Tests to Determine Ability

GM -73	MC -46	GC -66	FC -50	SG -64	EB -46
HT -56	OR -40	RC -53	GS -53	HF -63	CM -56
FB -46	ES -33	EB -46	TS -46	FK -60	DS -43
SH -60	HS -46	LR -53	HF -33	NP -50	BK -31

Average Physics Grades of the Above Students

GM -88	GC -94	SG -86
HT -84	RC -90	HF -85
FB -90	EB -85	FK -80
SH -75	LR -83	NP -83
MC -84	FC -80	EB -80
OR -72	GS -78	CM -70
ES -85	TS -83	DS -80
HS -78	HF -81	BK -79

When the physics grades and the preliminary test marks were averaged we have the following result. The students were divided into groups of eight each so that the total values of the three groups were approximately equal, viz:

	Group A	Group B	Group C
	GM -80	GC -80	SG -75
	HT -70	RC -71	HF -74
	FB -68	EB -66	FK -70
	SH -67	LR -68	NP -66
	Mc -65	FC -65	EB -63
	OR -56	GS -56	CM -63
	ES -59	TS -65	DS -62
	HS -62	HF -57	BK -54
Total	527	528	528

Result of the preliminary test for standardization of the three lessons:

Pupils	Lesson I.	Lesson II.	Lesson III.
A	25	15	18
B	14	8	11
C	13	18	12
D	6	19	19
E	9	15	19
Total	67	75	79

On the basis of this test the lessons were revised so the first was made a little more difficult and the last easier.

6. Final Experiments.

Tables I, II and III give a complete record of the various tests in the order of their sequence of performance.

TABLE I.—LESSON I.

Group A.—Text-book	mmediate	1 Week	4 Weeks
GM	67.5	77.5	67.5
HT	75	67.5	62.5
FB	70	55	45
SH	57.5	55	45
MC	90	80	50
OR	62.5	75	45
ES	80	7.5	60
HS	85	75	60
Total	587.5	562.5	435
Group B—Lecture			
GC	85	70	62.5
RC	62.5	55	45
EB	70	45	32.5
LP	90	72.5	50
FC	52.5	55	55
GS	70	40	35
TS	80	60	50
HF	45	55	40
Total	555	452.5	370
Group C.—Laboratory			
SG	75	55	55
HF	70	60	60
FK	70	50	62.5
NP	75	52.5	55
EB	70	55	55
CM	80	62.5	55
DS	90	67.5	65
BK	70	60	57.5
Total	600	462.5	465

TABLE II—LESSON II.

Group A.—Laboratory

	Immediate	1 Week	4 Weeks
GM	87.5	85	70
HT	77.5	67.5	45
FB	55	50	50
SH	70.5	55	50
MC	75	50	50
OR	70	50	40
ES	90	55	55
HS	70	65	60
Total	595	477.5	420

Group B.—Text-Book Recitation.

GC	85	85	45
RC	70	75	35
EB	50	45	10
LP	67.5	30	25
FC	80	67.5	55
GS	60	35	25
TS	70	72.5	60
HF	60	40	25
Total	542.5	450	285

Group C.—Lecture.

SG	75	55	37.5
HF	87.5	45	35
FK	75	70	50
NP	70	53	35
EB	65	45	60
CM	77.5	45	45
DS	90	80	85
BK	60	50	55
Total	600	445	402.5

TABLE III.—LESSON III.

Group A.—Lecture.

	Immediate	1 Week	4 Weeks
GM	70	60	45
HT	58.5	45	40
FB	55	25	40
SH	60	50	45
MC	35	35	35
OR	60	40	40
ES	75	45	55
HS	52.5	45	30
Total	466	345	330

Group B.—Laboratory

	Immediate	1 Week	4 Weeks
GC	55	40	32.5
RC	65	52.5	40
EB	60	42.5	35
LP	50	40	35
FC	60	62.5	30
GS	30	25	30
TS	70	50	30
HF	45	37.5	25
Total	435	350	257.5

Group C. Text Book.

	Immediate	1 Week	4 Weeks
SG	67.5	45	50
HF	50	40	35
FK	65	70	40
NP	65	55	50
EB	77.5	55	55
CM	70	60	50
CM	70	60	50
DS	80	70	70
BK	45	45	37.5
Total	520	440	387.5

In comparing the results of Table I we find that Group C taught by the laboratory method has the highest record, for the immediate test and for the final test. The one week test favors the text book method. Group B taught by the lecture method is the poorest for all three tests.

Table II shows the results of Lesson II, "Why is Baking Powder used in cake?" and indicates approximately the same standing for each group as in Table I. Group C shows the highest record for the immediate test. This group was taught by the lecture method. However for the one week and the four week test Group A, taught by the laboratory method, give the best results. Thus in both Lessons I and II the laboratory gives the best permanent results. The immediate result varies, in the first case it is the laboratory and in the second case it is the lecture method which is the best. In each case, however, it is Group C that receives the highest immediate result.

In Table III we find that Group C again receives the highest score, and this time it is with the text-book method. Both the one-week and the four-week tests for retention are like the immediate reproduction and Group C or the text-book method gets the honor. The laboratory method here is the poorest for both the immediate and the final tests. Group C in every case has the highest result irrespective of the method, which indicates that this group is slightly superior in mental ability to the other two groups. It would then be a better test of the methods to compare the results of each group in regard to the three lessons taught that group. In other words compare the lessons rather than the groups. This will eliminate the difference in mental calibre of the groups but will not consider the difference in difficulty of the lessons.

The results compared according to lessons are:

TABLE IV.—*Group A.*

Lessons	Immediate	1 Week	4 Weeks
Lecture–3	466	345	330
Text-Book–1	587.5	562.5	435
Laboratory–2	595	477	420

Here we see that the results again differ. In the immediate test the laboratory method is the highest but the retention is best by means of the text-book method. The lecture method in all three tests is far inferior to the other two, the text-book and the laboratory.

TABLE V.—*Group B.*

Lessons	Immediate	1 Week	4 Weeks
Lecture–1	555	452.5	370
Text-Book–2	542.5	450	285
Laboratory–3	435	350	257.5

Here the lecture method gives the best results both immediate and delayed. The text is second and the laboratory last. The results show very little difference in the percent. of retention in the different lessons.

TABLE VI.—*Group C.*

Lessons	Immediate	1 Week	4 Weeks
Lecture–2	600	445	402.5
Text-Book–3	520	440	387.5
Laboratory–1	600	462.5	465

In the last group, the lecture and the laboratory tie for first place, while after one week and four weeks the laboratory is conspiciously superior.

In all these tests there seem to be little superiority of one method over the other. This was also shown when we considered the work of the different groups in regard to the lessons. Even though the groups and lessons were of different calibre, the predominance of any one method would be apparent. This is not the case, for there is no record for any method that is conspicuously high in more than one or two cases.

We might look at the result now in toto for each group of tests according to the method taught.

The following tables give the results for the immediate, one week and four week tests in order to show the value of the method and also to check the lessons.

TABLE VII.

Immediate Resulty

Method	Lesson I.	Lesson II.	Lesson III.	Total
Lecture	555	600	466	1621
Text-book	587.5	542.5	520	1650
Laboratory	600	595	435	1630
Total	1742.5	1737.5	1421.0	4901

TABLE VIII.

Results After One Week.

Method	Lesson I.	Lesson II	Lesson III.	Total
Lecture	452.5	445	345	1242.5
Text-Book	562.5	450	440	1452.5
Laboratory	462.5	477.5	350	1290
Total	1477.5	1372.5	1135	3985

TABLE IX.

Results After Four Weeks.

Method	Lesson I.	Lesson II.	Lesson III.	Total
Lecture	370	402.5	330	1102.5
Text-Book	435	285	387.5	1107.5
Laboratory	465	420	257.5	1142.5
Total	1270	1107.5	975	3352.5

These tables show more distinctly the value of the three methods for giving immediate reproduction and delayed reproduction of facts. For immediate results the order is text-book best, laboratory second and lecture third. For results after one week we find the following order: Text-book, laboratory, lecture. And, for retention at the end of four weeks we find the following order: Laboratory, text-book, lecture. In every case the lecture proves the poorest, while the difference in the results of the laboratory and text-book methods is not so marked. From a study of the rate of forgetting it would appear that if we carried out these tests at still longer intervals the laboratory method might prove distinctly superior to the other two methods. The text-book method is shown by these studies to be superior to the other methods, but if we could extend the curve, we would perhaps find that the pupils taught by the laboratory method would retain the facts the longest.

We also see by the tables that the three lessons differ considerably in difficulty. The first two lessons are quite evenly matched, but the third lesson is appreciably more difficult. This is apparent in all three tests. The individual results for each lesson would therefore vary, and the truer record is obtained not by considering the results individually but as a whole as we have just done. We might even consider a complete total for the three methods, as follows:

TABLE X.

Tests	Text-Book	Lecture	Laboratory
Immediate	1650	1621	1630
One Week	1452.5	1242.5	1289.5
Four Weeks	1107.5	1102.5	1142.5
Total	4210	3966	4062
Average	1403	1322	1354

In order to compare the individual differences of the pupils for each method we may add the immediate and delayed results together and plot these. Such a graph will not only show the individual differences but also the relative value of each method. The data is given in Table XI.

TABLE XI.

Totals for Immediate and Delayed Reproduction.

	Pupil	Text-Book	Laboratory	Lecture
Group A	HS	220	195	127.5
	MC	220	175	105
	ES	217.5	200	175
	GM	212.5	242.5	175
	HT	205	190	143.5
	OR	182.5	160	140
	FB	170	155	120
	SH	157.5	175	155
Group B	GC	215	127.5	217.5
	TS	202.5	150	190
	FC	202.5	152.5	162.5
	RC	180	157.5	162.5
	HF	125	107.5	140
	LP	122.5	125	212.5
	GS	120	85	145
	EB	105	137.5	147.5
Group C	DS	220	222.5	255
	EB	187.5	180	170
	CM	180	197.5	167.5
	FK	175	182.5	195
	NP	170	182.5	158
	SG	162.5	185	167.5
	BK	127.5	187.5	165
	HT	125	190	167.5
Average		175	169	165
M. V.		30	28	23

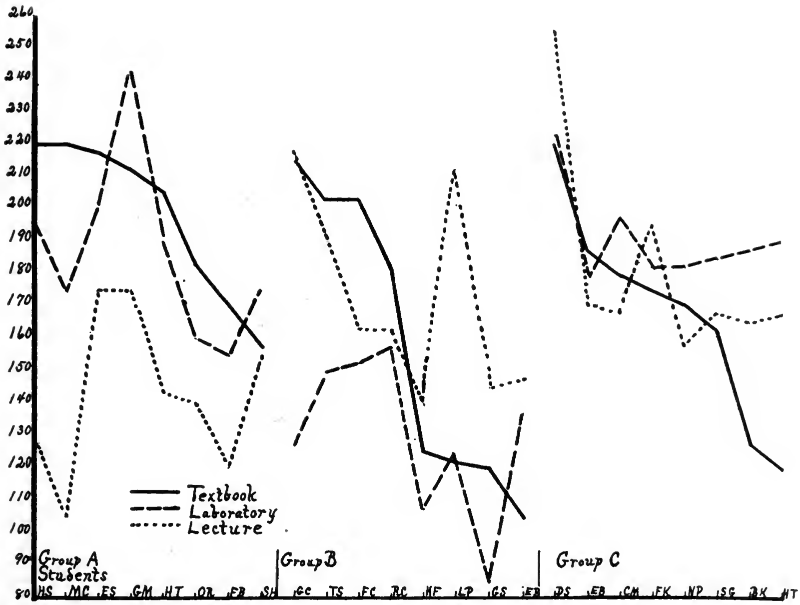


Figure 1. Total scores made by individuals of each group by each method.

. Figure I shows the individual variation in the different methods. The curves in the first part, Group A, are quite regular and become very irregular in the middle part, Group B. In the last part, or Group C, they follow each other quite regularly again. The high record held by the text-book method is quite apparent all through the graph. In the first part the lecture method is lowest, in the second part the highest, and in the third part it holds second place. In the third part of the graph the laboratory method holds the first place. This graph shows better than any of the others the great variations, not only in the groups, but among the individuals themselves. Inspection of Table XI shows that the larger the average the greater is the mean variation for the three methods.

So far we have considered only the quality of the work done, irrespective of the time taken to accomplish it. But economy in learning must consider two factors; the quality of the results, and the time required to accomplish the results. Equal results may be produced by any method, but economy in learning requires that we get the result with the least expenditure of time. The time required to teach each lesson may be tabulated thus:

TABLE XII.

Lessons	Text-Book	Lecture	Laboratory
I	28	22	53
II	21	33	38
II	30	28	41
Total	79	83	132
Average	26.3	27.7	44

The total results of the methods are:

TABLE XIII.

Lessons	Immediate	One Week	Four Weeks	Time
Lecture	1621	1242	1102.5	27.7'
Text-Book	1650	1452.5	1107.5	26.3'
Laboratory	1630	1290	1142.5	44'

We find that for immediate reproduction the text-book method yields the best results and takes the least time and is therefore the most economical. However for permanent results the laboratory method yields slightly the best results and the amount forgotten is less, but the time is far greater than by either of the other methods. If it is merely a question of immediate result the laboratory method is wasteful, and even when we consider the amount of retention after four weeks the slight superiority of the laboratory method scarcely compensates for the extra expenditure of time. There are other advantages, however, that the laboratory method has over the text and lecture methods which we cannot consider here.

SEX DIFFERENCES

On the basis of the material which we have at hand we may inquire what differences are shown by boys and girls in connection with the different methods of teaching chemistry.

TABLE XIV.

Average Scores by Sexes.

Method	Immediate		One Week		Four Weeks		Total	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Lecture	67.5	69.4	52.5	51.3	45.6	46.1	165.8	166.8
Text-Book	68.4	69.4	63.7	59.0	40.3	49.7	172.5	178.2
Laboratory	60.3	71.7	50.9	55.2	43.7	49.5	155.0	176.4

The average results of the different methods considered on the basis of sex show that in every method the girls outstrip the boys. In the immediate results the girls are higher in every method and exceptionally so in the laboratory method. The same is true of the final results but here the difference is great in both the text and laboratory methods. In the lecture method the difference is much less. The superiority of girls in retention after longer periods may be due to the nature of the lessons selected.

The distribution of grades, shown in Table XI and graphically in Figure 1, indicates that the tests were unequally balanced. As was pointed out in connection with Tables VII-IX, the third lesson was distinctly more difficult than the other two. This third lesson was given to Group A by the lecture method, to Group B by the laboratory method, and to Group C by the text-book method. Hence these three curves in the corresponding groups are unduly low.

From the standpoint of hygiene, there is no doubt as to the superiority of the laboratory method. From observation of the pupils during the teaching of the lessons their attention and interest reached its climax in the laboratory work. Restlessness, sliding down in the chair, looking out of the window, looking around the room, tapping with a pencil, with the fingers and feet, absent-mindedness, etc., were all noticeable in the lecture method especially. In the text-book method some such signs were apparent but not nearly so much as in the lecture method. Attention and interest, however, in the laboratory method seemed always at a high point. I know of no instant when the students were not absorbed in the laboratory work. As far as fatigue is concerned the periods were too short to produce much of it with pupils of this age. Early signs of it were perceptible, however, in the lecture method, but not in the laboratory work. The laboratory method is therefore the most hygienic and the lecture method the least.

Conclusion

1. There is not as great difference as is ordinarily supposed in the value of the three methods, lecture, text-book and laboratory, so far as imparting knowledge is concerned.
2. For immediate learning the text-book method is unquestionably superior.
3. For permanent learning the laboratory method is perhaps slightly superior.

4. In every respect the lecture method is the least effective in imparting knowledge to high school students.

5. From the point of view of the expenditure of time the laboratory method is the most costly and the textbook method the least. Ranking the methods in the order of economy of time they would stand: (1) Text-book, (2) lecture, (3) laboratory.

6. The rate of forgetting is greatest with the text-book method and least with the laboratory method. Whether the difference is enough to counterbalance the greater expenditure of time and thus make the laboratory method the most efficient is open to question.

7. Girls learn more readily than boys by all the methods, and more especially by the laboratory method. The question might be raised whether this is due to a greater retentiveness on the part of the girls, or a more careful attention to details in the learning process.

8. The different methods show decided individual differences both for immediate and delayed reproduction.

9. Probably a combination of the three methods will give the best results in the teaching of high school chemistry.

A TEST IN FIRST YEAR CHEMISTRY[1]

J. CARLETON BELL
Brooklyn Training School for Teachers

In the development of tests of intelligence a distinction has been drawn between knowledge tests and performance tests. One of the chief criticisms that have been made of the tests in the Binet scale is that they are too specific, that they call for relatively isolated bits of knowledge, that they are more indicative of the range of acquired information than of native abilities. Therefore we find other investigators, as Goddard, Healy, Terman, and Pintner, turning their attention to performance tests as more symptomatic of basic intelligence.

In the experimental investigation of attainments in school subjects a similar distinction may be made between tests of the range of information and tests of skill in performance. Here the performance tests have received by far the most attention. For this there are at least two reasons. First, the greater part of the elementary school course is devoted to formal, instrumental, rather than content subjects. The work in reading, in arithmetic, in drawing, in handwriting, and to some extent in spelling and grammar has been directed chiefly to the development of a generalized ability to meet situations rather than to the acquisition of specific informational content. That such ability is valuable and necessary is not to be denied, but there is reason to believe that the formal, drill aspects of elementary school work have been over-emphasized, to the neglect of knowledge of physical and human nature, history, literature and art. The same is true in large measure of the high school. Second, it is easier to measure such assumed general ability than it is specific information. All of our reading scales are constructed on the basis of testing the general ability to read, with little regard to the content of what is read. We have no scales to measure the child's knowledge and appreciation of poetry, of biography, of history, of novels, of short stories. All work in elementary arithmetic is necessarily formal, but until recently we have confined our attention to the four fundamental operations and a few random problems in reasoning. Few attempt have been made to explore the child's knowledge of fractions, of decimals, of denominate numbers,

[1] A preliminary account of this study was read before the Chemistry Section of the Central Association of Teachers of School Science and Mathematics at Columbus, Ohio, December 1, 1917, and appears in the May number of *School Science and Mathematics*.

of percentage, of involution and evolution. No scientific studies have yet been made of the school child's knowledge of animal and plant life, of the structure and functioning of his own organism, of his relations with his social environment, and we have only a few fragmentary efforts to determine his knowledge of geography and history. The range of possible knowledge in these subjects is so great and the attainments of a given child depend so much upon the circumstances of his environment that it is extremely difficult to construct tests that will give an adequate measure of the range of his information.

In attempting to determine the attainments of high school pupils in a natural science, like chemistry, we are again confronted by the performance and the informational aspects of the situation. We may ask what can the child *do* in chemistry, or we may inquire what does he *know* about the facts of chemistry. A complete measure of his attainments would, of course, include both. In determining his performance we might take him to a chemistry laboratory and confront him with a series of chemical situations typical of the work in elementary chemistry. This would be the most concrete and thorough test of his performances in chemistry, but would involve much time, expense, and effort in carrying out the test. A second method of procedure would be to confront him with a list of problems and ask him to indicate on paper what he would do in the solution of these problems. This type of test lacks the concreteness of the actual laboratory situation, but requires less time and expense, can be conducted with a large group of students at one time, and has the added advantage of affording a test of the pupil's scientific imagination. The writer believes that this type of problem test offers a rich field for experimental investigation which has been almost entirely neglected.

The present study was confined to the informational aspect of elementary chemistry. The question proposed was, what do high school pupils who have had one year of chemistry know about the subject? The study was further limited geographically to certain representative high schools in the state of Texas. The first problem in such a study is to determine the scope of elementary chemistry and the relative importance of the main topics taken up in the course. The ideal method of procedure here would have been to ascertain from the teacher of chemistry in each of the schools in question the topics covered in the course, and the appropriate amount

of time and attention devoted to each. This would have taken more time than could be allowed for the preliminary work, so recourse was had to a study of representative texts in elementary chemistry and of the recommendations sent to high schools by the chemistry department of the University of Texas. These recommendations are generally followed in the teaching of chemistry in Texas high schools, and serve to standardize the treatment of the subject. The next task was the selection of the specific questions that should constitute the test. If time had allowed, it would have been desirable to make an extensive preliminary study here by formulating a considerable number of questions on each topic, perhaps 200 or 300 in all, and submitting these to a large number of representative teachers of chemistry in different parts of the country (or perhaps merely to the teachers of the schools to be studied) with the request that they either rank the questions in the order of their importance, or indicate in some other manner the ones they considered most important. From a statistical treatment of the returns one could readily determine the questions that were considered of greatest importance by teachers of chemistry, and on this basis one could construct a test that would cover the most important points in elementary chemistry. In the present instance such an elaborate study was inexpedient, so that the questions had to be selected on the basis of individual opinion, supported by personal consultation with as many teachers of chemistry as could be interviewed.[2]

In formulating these questions the effort was made to state the question briefly, clearly and unambiguously, to call for such information as could be given in a few words, and to so frame the question that the answer could be marked as either right or wrong. All answers were to be written on the question sheet. When the papers were returned, however, it was found that some of the answers were partly right and partly wrong, so it was decided that a better indication of the abilities of the pupils would be afforded by giving one credit for correct answers, one-half credit for answers partly correct, and no credit for those entirely wrong. At the top of the sheet were blank spaces in which the pupil was asked to write his name, the name of the school, the date, the amount of chemistry he had studied, and the date of this study. The questions proposed were as follows:

[2] For this work the writer desires to express his obligation to Mr. R. G. Upton, an experienced teacher of chemistry, who was at that time assistant in the school of the Art of Teaching at the University of Texas.

CHEMISTRY TEST

1. What chemicals are liberated in electrolysis of water?
2. What gas is given off by the action of yeast in bread dough?
3. Name two chemicals used in making oxygen.
4. Express in cubic centimeters one litre.
5. Name two substances used in making hydrogen.
6. Define oxidation.
7. Define reduction.
8. Name four processes of purifying water.
9. What is a deliquescent substance?
10. Name two gases either of which might be used to bleach cloth or flowers.
11. Name two very common organic acids.
12. Name two chemical compounds used in making common salt.
13. Write the reaction for sulphuric acid on potassium hydroxide.
14. Name the two most abundant elements in the atmosphere.
15. Name two commercial sources of ammonia.
16. Is air a chemical or physical mixture?
17. Why does ammonia water always feel cold when it comes in contact with the hands?
18. Name substances used to make aqua regia.
19. Name four basic hydroxides.
20. Name the commercial uses of the nitrates.
21. What are the properties of carbon bisulphide?
22. Give chemical name of $CaSO_4$.............., H_2S................, Na_2SO_3...............
23. What is the valence of H............, of O............, of Ca.............., of SO_4................
24. Name the members of the chlorine group.
25. How many grams of water are formed by the combustion of 10 grams of hydrogen in the air? (Work on this paper).

Returns were received from the chemistry classes in the high schools of sixteen Texas cities. The city of Galveston returned papers from an elementary class and from an advanced class, and if these are counted separately (as is done in the tables) we have data from seventeen groups. Texas high schools begin their year's work about September 15. The papers were sent to the schools about May 10, and the tests were given about May 15. As the schools close about June 1, the work of the year had been practically completed. The pupils in fourteen of these schools had thus studied chemistry just eight months. In San Antonio the work had been begun late in the year, so that three months' study was reported. Thus San Antonio and the two groups from Galveston (one with three months' study and one with twelve months', should be considered by themselves.

The Results

A summary of the results is presented in Table I. The cities in which the high schools are located are given in alphabetical order. Column one gives the number of pupils in each group. It is probable that in the larger cities, as Fort Worth, Dallas, and Corpus Christi, these groups were divided into two or more sections, but there was nothing in the returns to indicate the sections. The classes are seen to be of moderate size, and are probably representative of chemistry classes in towns of less than 100,000 inhabitants. Each group is composed of both boys and girls, and of the entire number of four hundred papers received, one hundred and ninety-eight were from boys and two hundred and two from girls. It is somewhat surprising that so much interest should be shown in chemistry by girls. It is ordinarily assumed that girls prefer such studies as languages, literature and history, and that in elective courses in natural science boys predominate. It would be interesting to know whether studies in other parts of the country would show a similar equality in numbers of boys and girls taking chemistry.

As the returns were received, the papers were carefully gone over and each answer was marked right, partly right, or wrong. These marks were then transferred to score sheets presenting the individual responses made by the pupils to each question. Each pupil's total score was then computed and expressed as a per cent. of the maximum. The average of these per cent. scores was taken as the index of the group. Likewise the number of correct responses made by the group to each question was computed and expressed in terms of per

TABLE I

Percentage Score on Each Question

	No. Pupils	Months	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	Average	No. Boys	Av. for Boys	No. Girls	Av. for Girls
			No. of Question																													
1. Beeville	23	8	87	91	83	9	76	58	48	35	30	72	83	35	61	65	24	48	22	67	67	24	24	70	78	83	26	56	9	54	14	57
2. Brownwood	17	8	100	100	100	83	100	95	89	75	64	83	95	75	75	92	50	95	42	86	97	64	31	61	89	100	39	78	11	75	6	76
3. Coleman	15	8	100	100	97	74	93	67	63	53	93	80	87	40	74	80	17	80	63	74	87	50	40	60	60	90	13	69	3	72	12	68
4. Corpus Christi	44	8	83	82	77	68	88	86	83	40	11	65	31	61	97	69	12	48	48	73	91	22	15	64	99	98	18	59	20	64	24	55
5. Dallas	46	8	100	98	62	42	96	89	84	77	39	76	51	46	64	84	34	70	28	58	87	40	24	85	86	57	63	63	16	76	30	56
6. Del Rio	13	8	100	100	69	0	69	69	77	65	69	81	92	69	89	81	8	69	58	81	69	8	15	77	85	92	27	65	8	61	5	71
7. Denton	18	8	95	95	36	67	97	33	33	67	28	64	19	61	75	75	6	95	39	11	58	19	22	47	70	92	6	52	14	54	4	44
8. El Paso	21	8	95	91	21	79	57	76	79	53	55	62	55	60	95	88	26	76	60	84	76	41	31	76	88	95	21	65	15	66	6	55
9. Ft. Worth	50	8	97	99	71	77	79	76	66	65	76	83	44	68	75	91	13	78	18	77	89	47	12	66	85	98	54	69	16	71	34	68
10. Garland	16	8	94	81	81	44	94	50	19	53	31	59	31	28	3	56	22	31	19	91	78	22	9	41	3	88	0	45	7	48	9	42
11. Lockhart	14	8	100	100	82	72	93	100	100	89	93	68	93	93	97	93	57	72	29	61	100	39	46	75	93	100	72	81	7	87	7	75
12. Mexia	21	8	95	95	50	86	62	55	19	53	53	74	0	17	50	57	2	67	33	33	53	26	0	57	36	91	10	47	12	53	9	39
13. San Marcos	28	8	70	29	7	32	29	39	7	41	0	57	43	21	57	54	2	39	29	9	86	9	34	57	61	63	0	36	13	38	15	34
14. Uvalde	10	8	90	100	0	80	50	60	60	45	60	65	30	20	30	70	0	60	0	40	45	55	0	50	45	40	40	45	6	44	4	47
Average	24		93	90	60	58	77	62	59	58	50	71	54	50	67	75	20	66	35	60	77	33	22	63	70	85	28	59	11	62	13	56
15. Galveston, Adv	14	12	100	93	79	72	89	100	86	72	68	72	79	79	97	86	39	86	43	82	89	57	18	86	86	100	36	76	8	79	6	73
16. Galveston, Elem	14	3	90	7	50	93	50	57	0	46	7	43	11	39	61	82	0	79	21	0	68	14	0	29	50	0	0	36	10	39	4	29
17. San Antonio	36	3	97	83	67	60	78	76	33	54	72	71	58	53	56	88	32	51	50	39	56	40	18	61	60	58	28	58	23	60	13	54

cent. of the maximum. These percentages for each question are shown in Table I in the columns numbered from 1 to 25. The column marked Average presents the average total score or index for each group. Examination of this column shows that the average index for the fourteen schools whose pupils had studied chemistry for eight months was 59. The range is from 81 (Lockhart) to 36 (San Marcos). This ratio is well over two to one, and shows that there is a vast difference in the chemistry teaching in Texas high schools. This sets an interesting problem which calls for further investigation. What is the reason for this wide range of scores? That it is due to individual differences in pupils is scarcely likely. One might think that in the larger schools the individual pupils do not get as much attention, and hence their showing is not so good. Inspection of the table, however, does not bear out this assumption. On the contrary, the larger schools, Fort Worth, Dallas, and Corpus Christi, show a combined score that is well above the average, while the small schools, Garland, Mexia, and Uvalde, are near the bottom of the list. In the absence of other evidence we suspect that the chief reason for the difference is to be found in the individual teacher, and that in general the larger towns pay the better salaries and get the better teachers. Lockhart, Brownwood and Coleman, however, all small towns, stand at the very top of the list, and thus tend to weaken the force of this hypothesis.

Table II gives the percentile distribution of the individual scores made in each school. It will be seen that the score of the best pupil is usually more than twice as good as that of the poorest in the same class, and in one case (San Marcos) it is almost four times as good. Marked exceptions to this are Brownwood and Lockhart, where the pupils are fairly close together, and all stand high. From the entire body of four hundred pupils, only one (Roland Bond, of Dallas) made a perfect score, or 100 per cent. Other high scores were made by Pliny Dell Valle, Dallas (98 per cent.); J. D. Miller, Dallas (92 per cent.); Alfred Harrys, Denton (94 per cent.); Hilmar Karbach, Lockhart (94 per cent); and Harry Galloway, Lockhart (92 per cent). On the other hand the lowest score (16 per cent.) was made by a boy in the San Marcos school, and of the lowest ten scores six are to be credited to that school. This difference between the highest and the lowest schools can be clearly seen in Table I. Lockhart, with an index of 81, had 100 per cent. correct on six questions, and 75 per cent. or better on sixteen of the twenty-five questions, while the

TABLE II

Range of Individual Scores in Each School

	Lowest	Percentile									Highest
		10	20	30	40	50	60	70	80	90	
1. Beeville	18	36	38	38	50	56	58	62	66	74	90
2. Brownwood	66	68	70	76	78	80	82	82	86	86	88
3. Coleman	34	50	60	62	70	74	74	80	80	84	88
4. Corpus Christi	42	48	50	54	56	60	64	66	72	78	86
5. Dallas	40	46	52	56	60	66	68	70	78	80	100
6. Del Rio	40	56	58	58	62	66	70	74	78	78	82
7. Denton	32	34	38	42	46	50	52	54	62	86	94
8. El Paso	46	46	52	60	62	68	70	72	78	78	84
9. Ft. Worth	38	58	62	66	68	70	72	74	78	80	84
10. Garland	24	34	36	38	40	44	46	46	52	60	70
11. Lockhart	66	72	76	78	80	82	84	84	88	90	94
12. Mexia	28	34	38	40	42	48	50	52	54	64	70
13. San Marcos	16	18	22	26	34	38	46	48	52	52	62
14. Uvalde	22	30	32	38	44	46	46	58	58	58	78
15. Galveston, Adv	50	62	68	70	72	84	84	84	84	86	88
16. Galveston, Elem	20	28	32	34	36	36	36	40	42	44	54
17. San Antonio	30	32	38	42	48	56	60	68	76	82	86

lowest question had 29 per cent. San Marcos, with an index of 36, made no higher score than 86 per cent. on any question, had only eight questions over 50 per cent., received scores of less than 10 per cent. in seven questions, and failed entirely on two questions. The widest difference between extreme scores is seen in Beeville, where the range is from 18 to 90.

In addition to the scores of the fourteen schools already discussed, Table I presents the results from the two classes in Galveston and from the San Antonio high school. Evidently the elementary class in Galveston had not gone very far in chemistry, for on six of the questions no score at all was made. The index for the class was 36, which although it is not high, is proportionately much better than the showing made by some of the schools after eight months of study. The advanced class, with twelve months of study to its credit, presents an index of 76, which is higher than that of any other school except two, but is carecely a showing adequate to the greater amount of time spent on the subject. The large class in the San Antonio

high school, on the other hand, after only three months of study, attains an index of 58, a score within one point of the average of the indices of all the eight months schools. Moreover, an inspection of the scores made on the individual questions shows that the attainment was remarkably even and well-balanced. With the time and other factors included in the reckoning perhaps San Antonio makes the best showing of any of the schools.

Attention has already been called to the fact that the number of boys and girls taking the test was practically equal. This does not apply to the individual schools. In the larger schools, as Fort Worth and Dallas, the girls outnumber the boys approximately two to one, while in the smaller towns the boys are usually in the majority. In the general average of attainments, however, the boys are distinctly superior to the girls, the scores standing 62 for boys to 56 for girls. This is not only true of the general average but applies to the schools individually, only four classes of the seventeen showing a higher score for the girls than for the boys. The superiority of the boys is especially marked in Corpus Christi, Dallas and Lockhart, while in Fort Worth and San Antonio it is not so pronounced. It is noteworthy that the group of pupils having the highest ten scores is composed exclusively of boys. Likewise the lowest two scores were made by boys, out of the lowest ten five were boys and five were girls. So far as it goes, therefore, this investigation supports the contention that boys are more variable than girls

Table I shows not only the average score of each school on each question, but also the average scores of the fourteen schools that had had chemistry for eight months. These averages are made directly from the school scores, and are, therefore, on the basis of the same number of pupils for each school. It is of interest to compare the different questions from the point of view of the responses made to each. Table III presents the questions in order of increasing difficulty, or at least in the decreasing magnitude of the average score. As might be expected, the electrolysis of water is the most familiar, with the action of yeast in bread dough a close second. These facts are easily retained, owing to their connection with daily life. The knowledge of the members of the chlorine group is more technical, and it is somewhat surprising that the question receives such a high score. These three questions are, therefore, in a group by themselves. The next seventeen questions show a regular decrease in score from 77 to 50. Seven of these questions lie within

TABLE III.

Questions Arranged in Order of Increasing Difficulty.

Per cent.	No. of Question.	Question.
93	1.	What chemicals are liberated in electrolysis of water?
90	2.	What gas is given off by the action of yeast in bread dough?
85	24.	Name the members of the chlorine group.
77	5.	Name two substances used in making hydrogen.
77	19.	Name four basic hydorxides.
75	14.	Name the two most abundant elements in the atmosphere.
71	10.	Name two gases either of which might be used to bleach cloth or flowers.
70	23.	What is the valence of H, of O, of Ca, of SO_4?
67	13.	Write the reaction for sulphuric acid on potassium hydroxide.
66	16.	Is air a chemical or physical mixture?
63	22.	Give the chemical name of $CaSO_4$, H_2S, Na_2SO_3.
62	6.	Define oxidation.
60	3.	Name two chemicals used in making oxygen.
60	18.	Name substances used to make aqua regia.
59	7.	Define reduction.
58	4.	Express in cubic centimeters one litre.
58	8.	Name four processes of purifying water.
54	11.	Name two very common organic acids.
50	9.	What is a deliquescent substance?
50	12.	Name two chemical compounds used in making common salt.
35	17.	Why does ammonia water always feel cold when it comes in contact with the hands?
33	20.	Name the commercial uses of nitrates.
28	25.	How many grams of water are formed by the combustion of 10 grams of hydrogen in the air?
22	21.	What are the properties of carbon bisulphide?
20	15.	Name two commercial sources of ammonia.

a range of five points (63 to 58). The last five questions are so much lower that they form another distinct group. Perhaps Question 17 deals with a phenomen that is more physiological than chemical, but the facts are usually brought out in connection with laboratory work. Evidently the commercial uses of nitrates and of ammonia receive very little attention in these chemistry classes. Is this another indication of abstract formalism in science teaching, and of neglect of the concrete applications of the science?

Table IV affords a comparison of the attainments of the best ten and the poorest ten of the entire four hundred pupils. Of the twenty easier questions the good group made 100 per cent. on fifteen, 95 per cent. on four, and 85 per cent. on one. Their lowest score was

TABLE IV.

Average of Highest Ten and Lowest Ten Pupils

	Average Total Sorce	No. of Question 1	2	3	4	5	6	7	8	9	10	11	12	13
Average of highest ten pupils..	.93	100	100	95	100	100	100	100	95	100	95	100	100	100
Average of lowest ten pupils..	..20	65	20	5	50	10	20	0	30	10	35	20	5	10

	14	15	16	17	18	19	20	21	22	23	24	25
Average of highest ten pupils.............	100	55	100	70	85	95	70	75	100	100	100	80
Average of lowest ten pupils..............	50	5	30	10	5	50	5	0	10	30	15	0

55 per cent., on the commercial uses of ammonia. The scores of the lowest ten seem to be quite erratic. They agree with the others in showing the greatest familiarity with the components of water, but otherwise there is no agreement either with the highest ten or with the entire group. On three questions they get 50 per cent. One of these is purely mathematical, one deals with the elements in the atmosphere, and only the last (name four basic hydroxides) depends strictly on a study of chemistry. Three of the questions (Nos. 7, 21, and 25) arouse no trace of accurate response.

The present study is tentative, incomplete, and limited in its scope. It would be desirable to have more information regarding the text-book used in each school, the topics emphasized in the course, the number of recitations per week, the amount and character of the laboratory work, and the point of view of the teacher. The range of questions might well be amplified, and correlated studies should be carried on with other high school subjects. A study of school marks in chemistry should go along with such testing. It is hoped that this article may stimulate others to make similar and more thorough investigations.

THE RANGE OF INFORMATION TEST IN BIOLOGY. I. PHYSIOLOGY*

N. M. GRIER

Central High School, St. Louis, Mo.

The fact that the human mechanism has been studied previously in the grades, suggests two possibilities to those concerned with the effective teaching of high school physiology and its proper place in the course of study:

1. The knowledge acquired in the grades may be disjointed and fragmentary. In the words of a writer on the relation of this subject to high school zoology: "We may not assume that the pupils have brought from the elementary schools any accurate knowledge of physiological principles and processes." If this be so, the problem before the high school is evidently that of utilizing somehow this previous knowledge, whatever the type of course to be given.

2. On the other hand, physiology may hold no justifiable place in the high school course of study, for, barring anatomical detail and technical terminology, grade-school pupils, when entering the high school, may already possess a *reasonable* amount of information concerning the human body. In the opinion of the writer previously quoted, "High School physiology adds nothing but details, except as it enters upon the fields of the high school sciences, thus producing a wasteful duplication of work."

So far as the present writer is aware, neither of the statements of the author cited rests on detailed evidence from which a conclusion one way or the other could be drawn. Recognition of the two possibilities has resulted in the proposal of various new types of courses, such as "Social Biology," "Biology of Man," etc., diagnosed to meet both of them. With the idea of securing more definite data concerning these problems, the writer, at the suggestion of Dr. E. L. Thorndike, Columbia University, has conducted tests with first-year high-school students after the plan described by Whipple in his *Range of Information Test*. It is hoped that the results obtained afford partial solutions.

---

*Read before Section L, A. A. A. S., January, 1918. Grateful acknowledgement for helpful criticism is due Mr. T. S. Henry, University of Illinois. Thanks are due to Mr. W. J. S. Bryan, Assistant Superintendent of Instruction, and Mr. C. B. Curtis, Principal, Central High School, St. Louis, for facilities offered in making this test.

The 100 test words were selected from the textbook used in the grades in the following manner. A list was first prepared of all those words pertaining more directly to physiology discovered in the spellers and readers previously used by the pupils in their grade-school work. A similar list was made from a number of the articles dealing with pure food and health to be found in the daily newspapers and more widely known monthly magazines. From these lists, 100 of the more frequently occurring words were checked in the glossary of the textbook as being (1) words in fairly common use, which, it was thought, cultural values to be derived from the science demanded that the pupils should know, (2) words a knowledge of which would indicate some familiarity with the care of the body, and the activities of its organs. Many of these terms, such as paralysis, ligament, etc., are now only semi-technical. An effort was made to distribute the words evenly among the divisions of the subject.

To make clearer the use of the test in this connection, the test blank is here reproduced with some desirable but slight modification. The pupils were asked to read the directions through twice before marking the words, their attention was called to the request for definitions as printed below the test words, and they were given all the time they needed. As a check on the marking of the pupils, the 100 words were defined by the examiner from the glossary of the grade-school textbook, and each pupil revised his own paper by placing a second series of marks after each word to indicate the manner in which he should have marked it. A comparison of the number of D's, E's, F's and N's of the first and second series will then show with a fair degree of accuracy the extent and nature of the error due to ignorance or misunderstanding of the real meanings. The results obtained may be found in the following tables classified as to dependence of range of information of physiology on:

1. Academic status.
2. Sex.
3. Academic status and sex.

By first-term students are meant those taking up the study of botany at the time. Second-term students are beginning physiology, having completed botany.

DIRECTIONS

NAME .... DATE ....

Below are 100 words which are largely technical and which are designed to test the extent of the knowledge of physiology you gained in the grade school. Consider each word carefully.

1. Place a D before the terms you can DEFINE as exactly as words are ordinarily defined in the dictionary.
2. Place an E before the terms you can explain to one not familiar with their meaning.
3. Place an F before the terms with which you are roughly familiar.
4. Place an N before the terms which are new to you.
5. At the bottom write out the definitions of the first 5 words you marked D, and the first 5 you marked E.
6. Count the number of D's, E's, F's, and N's, and record the result at the top of the page in the 1 in. space.

D.	E.	F.	N.
1.	1.	1.	1. ....
2.	2.	2.	2. ....

abdomen
absorption
adulteration
antitoxin
aorta
artery
auditory canal
auricles
biceps
bicuspids
bile
bronchial tubes
bowels
capillary
carbohydrate
carbonic acid gas
cartilage
cerebellum
cerebrum
circulation of blood
congestion
convolution
cornea
corpuscles
contamination
cranium
curvature of spine
cuticle
diaphragm
digestion
disinfectant
dislocation
enamel
energy
epiglottis
epidemic
excretion
fainting
far-sighted
femur
forearm
fracture
germ
gland
gray matter
gullet
gymnastics
habit
hinge-joint
immunity
infection
inflammation
intestine
larynx
kidney
involuntary muscle
lead poisoning
ligaments
liver
lymph
marrow
massage
mastication
mucus membrane
muscle fiber
narcotic
nerve fiber
nitrogen
oxidize
patent medicine
pancreas
paralysis
periosteum
pigment
pneumonia
pore
pelvis
proteid
pulse
pupil
preservative
pus
quarantine
retina
saliva
sewer gas
spinal cord
sprain
stimulant
stomach
tartar
tendon
tissue
triceps
tuberculosis
urea
vein
ventricle
vocal cords
wind pipe

D.	E.
1.	1. ....
2.	2. ....
3.	3. ....
4.	4. ....
5.	5. ....

TABLE I

Dependence of Range of Information of Academic Status

A. (1st series)

ACADEMIC STATUS	NUMBER	D.	E.	F.	N.
1st-term pupils	40	17.3	13.3	18.1	51.3
2nd-term pupils	97	14.7	17.9	24.8	42.7

B. (2nd series)

ACADEMIC STATUS	NUMBER	D.	E.	F.	N.
1st-term pupils	40	24.9	17.2	11.5	46.4
2nd term pupils	97	20.3	24.1	22.1	33.4

Computation made from the first and second series of the preceding table indicates that with increase of maturity, there occurs also an increase in the number of technical terms that can be explained (E), or that are least familiar (F), a decrease in the number of terms that can be defined (D), among second-term students. The only reason for the latter which the writer may present is that the comparatively long time elapsing before high school physiology was taken up—a term and the summer vacation—had somewhat altered the conceptions of these pupils.

The increase of E's follows probably because certain of the test words are used in the preceding course, botany. Another factor possibly affecting the results is the knowledge obtained through the widespread propaganda against disease and for better living, inaugurated by various civic organizations, and the articles in magazines and newspapers forming a partial basis for this test. Knowledge of the body, of course, increases with age, and in a trial test of 12 young men of an average age of 20 (most of whom had never attended the high school), very high percentages of E's and F's were obtained and practically negligible percentages of D's and N's. The advantages of a high school course in physiology may be doubted under these conditions, a conclusion supported by the fact that in Table I-B, the number of words defined (D), explained (E), familiar (F), is decidedly greater than the words which are new. The average error of marking (difference between first and second gradings), in first-term students was 72 per cent., that of second term students, 64 per cent., showing that there also occurs with maturity some advance in discriminatory power.

These results, reclassified by sex, show a general superiority in boys in range of information of physiology. This is in accord with the results obtained by Whipple and Smith in their studies. Further evidence may be obtained for clearer concepts in boys than in

girls from the fact that the sum of N's in the second series is less and there is a greater percentage of D's and E's among the boys. The average error in boys was 56 per cent.; that of girls 8 per cent.

Attention may here be called to the fact that in most of the tables presented, there is a tendency for a greater percentage of D's to be that the percentage of error obtained by comparing these as groups is practically the same. As might be expected, the checking-up process diminishes the percentage of F's and N's, and increases the percentage of D's and E's.

TABLE II

Dependence of Range of Information on Sex

A. (1st series)

Academic Status	Number	D.	E.	F.	N.
High-school Boys	65	18.0	20.6	21.5	39.9
High-school Girls	72	13.1	12.9	24.0	50.0

B. (2nd series)

Academic Status	Number	D.	E.	F.	N.
High-school Boys	65	22.0	23.6	19.3	35.1
High-school Girls	72	21.4	21.0	18.5	39.1

The insight afforded by a mixed classification by sex and academic status is presented in Table III (A, B, C, D.) First-term boys (C and D) have a greater proportion of D's than first-term girls, and a slightly less number of E's. No difference with regard to F's and N's is noted from that of the preceding table. First-term girls however, have a lower percentage of error, 6.6 per cent., against 7.5 per cent.

Table III, (A and B), representing results from second-term boys and girls indicates that the boys secured a greater percentage of D's in the first series but that these percentages are practically equivalent in the second. In both series they had a greater percentage of E's and fewer F's than the girls. N's are distributed as in the preceding table. The percentage of error in the girls was 9.4 per cent.; that of the boys, 3.6 per cent. This may indicate a rapid change in mental powers at this period.

TABLE III

Dependence of Range of Information on Sex and Academic Status

A. (1st series)

Academic Status	Number				
2nd-term High-school Boys	55	16.8	22.2	22.5	38.5
2nd-term High-school Girls	42	11.8	12.2	27.8	48.2

B. (2nd series)

Academic Status	Number	D.	E.	F.	N.
2nd-term High-school Boys	55	20.8	25.4	20.6	33.2
2nd-term High-school Girls	42	19.9	22.9	23.6	33.6

TABLE III

Dependence of Range of Information on Sex and Academic Status

C. (1st series)

Academic Status	Number	D.	E.	F.	N.
1st-term High-school Boys	10	24.8	11.8	15.8	47.6
1st-term High-school Girls	30	14.8	13.8	18.8	52.6

D. (2nd series)

Academic Status	Number	D.	E.	F.	N.
1st-term High-school Boys	10	28.9	13.8	12.0	45.3
1st-term High-school Girls	30	23.5	18.3	11.4	46.8

A list of typical errors of definition follows—the word apparently the source of confusion indicated in parenthesis. Many illustrate erroneous impressions commonly met with in pupils, points which may furnish opportunities for objective attacks by teachers in solving the first problem indicated in this paper.

abdomen—the stomach: digestive part of body: bottom part of stomach containing the bowels.
absorption—taking up of water by air: taking in water or sunlight: taking in of oxygen.
adulteration—making food, drink or medicine weaker by adding water: to dilute with water.
antitoxin—a serum infected to certain diseases; mixture of blood of sick person to cure somebody else having same disease.
aorta—chief vein of body: blood vessel passing through lungs.
artery—a main vein: tube or muscle in the body carrying blood toward the heart: blood vessel for pure blood.
biceps—name applied to teeth inside of mouth (bicuspid): animal having two feet (biped): your legs.
bile—waste issuing from body after food is digested: place where food goes through (bowel).
bronchial tubes—tubes which make voice in throat.
bowels—part of stomach: part of stomach secreting waste: where food circulates.
carbohydrate—mixture of starch and sugar: a food consisting of nitrogen, sulphur and phosphate.
carbonic acid gas—gas obtained by eating zinc with hydrochloric acid (hydrogen).
cartilage—matter found in spinal column.
curvature of spine—caused by killing end of spine.
cerebrum—part of spine.
circulation of blood—way blood helps itself to flow: flow of blood through veins: working of blood in system.
congestion—to be crowded on outside: blocked passage-way.
convolution—cramps (convulsion): kind of nerve is drawn up and hit.
contamination—to catch a disease: disease that spreads.
cranium—term sometimes used in speaking of head.
cuticle—end of loose skin at finger nails.

diaphragm—line of dividing chest: separates stomach and bowels.
digestion—when the food is taken into the stomach: assimilation of food: process of chewing things and swallowing them: work of stomach preparing food for distribution over entire body: if the bowels move.
disinfectant—a liquid used for various cleaning purposes: exterminator of insects and germs: poison medicine: a substance to take away a bad odor.
dislocation—sprain or broken bone: breaking and changing a bone.
enamel—a kind of paint: white shiny layer of polish something of a granite color only white instead of spotted.
energy—will power: putting life into things.
epiglottis—outside skin (epidermis).
epidemic—drug used before and after operations to quite pain (anaesthetic). catching sickness in which 1-10 of population dies: over 100 cases of disease.
excretion—taking in of water by plants: overdoing it (exertion):
fainting—act of being unconscious: when heart loses 1 beat: weakening of nerves: falling into a certain mood: rushing of blood to top of your head.
far-sighted—cannot see but when he is far away: when one cannot see far: one who sees objects in front of them.
fracture—part of body dislocated.
germ—insect that carries disease: harmful insect: an insect in the body causing disease: a small insect that can be seen through the microscope.
gland—something that makes you see, as eye gland: part of neck.
gray matter—a poison formed when a cut is not attended to.
inflammation—you have a spasm or a fit (convulsion): when a person has a high degree of fever they are inflamed.
intestine—last stage of digestion: your inside consisting of kidney, bowels, lungs.
kidney—water preserver: an organ of the stomach.
ligaments—muscle tissues: kind of muscle in leg or arms.
liver—filters blood: part of the stomach.
marrow—white substance in the bone keeping them hard: fluid in bones.
oxidize—where carbon and oxygen unite to form a substance called oxide.
patent medicine—medicine for a special cause.
paralysis—drying up of marrow in bone but mostly in spinal cord: lameness of body which may cause death: stiffening of joints.
pupil—white part around part you see with: colored part of eye.
quarantine—to be kept in.
saliva—white matter from the mouth.
spinal cord—a continuous bone supporting the back: composed of 24 small bones from head to coccyx. (spinal column).
sprain—occurs when a bone is out of position: where not able to use ankle or arm.
stimulant—awake or moving: reviver.
tendon—something on the order of a ligament found in the legs.
tissue—soft meat of the body.
urea—salt water from body (urine).
vein—tube in body carrying blood to different parts of body.
ventricle—pipe in throat food goes down: sound without moving lips.
windpipe—tube which carries food to the stomach.

COMMUNICATIONS AND DISCUSSIONS

SCHOLARSHIP AND SUCCESS IN TEACHING

A question of perennial interest and one that provokes considerable discussion among educators and business men, "Is there any relation between scholarship and vocational success?" is the subject of this inquiry. The purpose of this study is to determine whether there is a relationship between scholarship and success in teaching. The data have been gathered from the rating in scholarship given students in the Harris Teachers College of St. Louis and the ratings given the same students as apprentices, temporary substitutes, permanent substitutes, and teachers in the St. Louis schools. The resident faculty members rate the students, and the grade school principals of the city rate the apprentices, the temporary substitutes, the permanent substitutes, and the grade teachers.

An explanation of terms may be necessary. An apprentice is a student who, after completing two terms of work in the College, spends a term in a grade school for observation and practice teaching, under direction. A temporary substitute is a graduate of the College who has not received permanent appointment and may be sent to any school where teachers are absent. The temporary substitute may go to a different school every day. A permanent substitute is a graduate who has not received an appointment as elementary school assistant, but who is assigned to a permanent place in the schools.

In a report of the Superintendent of the St. Louis schools, for 1911 and 1912, Dr. John W. Withers, then President of the Harris Teachers College, says: "A careful study of statistics of failures that occurred in all classes since the College began showed that almost all the failures were made either by those who came from other sources than the St. Louis public high schools, or by those whose record in these high schools placed them in the lowest third of the classes with whom they graduated. For this reason it was decided by act of the Board on May 14th, 1912, to admit to the College without examination, hereafter, only those graduates of the St. Louis public high schools whose record for the high school course placed them within the upper two-thirds of the classes in which they graduated, and to require that all other graduates of our own or other secondary schools should pass an entrance examination covering four years of high school work."

The object of the examination given for entrance to the college is to exclude all persons from the College who do not have scholarship equivalent to that possessed by those in the upper two-thirds of the classes in the St. Louis high schools, and this regulation of the Board of Education was passed because it had been determined that low scholarship in the high school meant low scholarship and failures in the College, and also failure in teaching. If the students, therefore, who have a low scholarship in the high school have a like record in the College, it remains to be determined whether the College and high school records correspond to those made when the student enters upon her teaching career. The following table shows the apprentice record as compared to the College record.

APPRENTICESHIP RECORDS OF 359 GRADUATES

	Excellent	Good	Medium
Management of Children.			
1st Third	65 %	32½%	2½%
2nd Third	42 %	54 %	4 %
3rd Third	25 %	62½%	12½%
Instruction.			
1st Third	65 %	34 %	1 %
2nd Third	46 %	52 %	2 %
3rd Third	19½%	67½%	13 %
Attention to Details of School Business.			
1st Third	92 %	8 %	0 %
2nd Third	86½%	12½%	1 %
3rd Third	57 %	38 %	5 %

"The table shows, for instance, that 65% of the highest third of the group, 42% of the middle group, and 25% of the lowest group won *e.'s* in the management of children. There is even a more marked correlation in the subject of instruction and a less marked, but still a positive, correlation in attention to details of school business."

To take the class of June, 1914, we find the following to be true: "The highest third of the class compared with the highest third of the six preceding classes won an average of 9% more *e.'s* in the management of children, 17% more *e.'a* in instruction, and 2% more *e.'s* in attention to the details of school business; the middle third won 5½% more *e.'s* in the management of children, 2 1-3% more *e.'s* in instruction, and 6 2-3% less *e.'s* in attention to details of school business than the corresponding third of the six preceding classes; the lowest third won 6% more *e.'s* in management, 21½% more *e.'s* in instruction, and 16% more *e.'s* in attention to details than the lowest third of the six preceding classes." Here, then, in the case of even a single class of 58 members we find an interesting, positive correlation between the scholarship record in the College and the judgment of the principals indicating the teaching efficiency of these young women.

But the reader may say that the apprentice record is not a fair comparison and that actual teaching efficiency can not be measured by scholarship rating in that way. Let him consult the following table which gives the record of two hundred ninety-five graduates as temporary substitutes, permanent substitutes, and as teachers.

Temporary Substitutes: 6 Classes—January, 1911, to June, 1913. 295 Graduates

	Management				Instruction				Attention to Details			
	%-E	%-G	%-M	%-U	%-E	%-G	%-M	%-U	%-E	%-G	%-M	%-U
H.3*	38.5	52.3	8.4	.8	29.8	63.9	6.0	.3	54.8	42.4	2.8	.0
M.3*	28.6	60.6	9.5	1.3	21.6	70.6	6.2	1.6	50.1	46.8	2.7	.4
L.3*	25.0	62.8	10.3	1.9	17.8	73.4	7.9	.9	45.9	50.4	3.3	.4

Permanent Substitutes: 6 Classes—295 Graduates

	Management				Instruction				Attention to Details			
	%-E	%-G	%-M	%-U	%-E	%-G	%-M	%-U	%-E	%-G	%-M	%-U
H.3	39.8	51.0	8.0	1.8	36.5	56.7	6.6	.2	64.9	32.7	2.3	.1
M.3	29.3	59.3	10.4	1.0	27.0	65.9	6.5	.6	59.9	37.7	2.4	.0
L.3	23.9	59.9	14.8	1.4	12.8	72.9	12.8	1.5	55.4	40.2	4.4	.0

Teachers: 3 Classes—144 Graduates

	Management				Instruction				Attention to Details			
	%-E	%-G	%-M	%-U	%-E	%-G	%-M	%-U	%-E	%-G	%-M	%-U
H.3	39.1	46.7	12.0	2.2	39.9	50.2	9.5	.4	69.8	28.2	2.0	.0
M.3	40.5	58.1	1.4	0.0	27.0	71.6	1.4	.0	79.9	17.6	3.5	.0
L.3	39.5	57.9	2.6	0.0	17.1	80.3	2.6	.0	78.6	20.0	1.4	.0

*H.3—Highest third: M.3—Middle third: L.3—Lowest third.

The classes are again divided into tertiles according to their relative rank at graduation and the percentage of *e.'s*, *g.'s*, *m.'s*, and *u.'s* won by each tertile in the management of children, instruction, and attention to the details of school business in their work as temporary substitutes, permanent substitutes, and teachers are presented in order. The table compares 17090 markings of principals, 10704 of which were on the service as temporary substitutes of 295 graduates, 5585 on the work of the same group as permanent substitutes, and 801 on the services of 144 graduates (Classes of January, 1911, and June, 1911, and January, 1912) as appointed teachers. The temporary substitutes' service covered a period of 12197 days, the permanent substitutes' 1975 months, and appointed teachers' 196 terms. It will be noted that the percentage of *e.'s* and *g.'s* is very high amounting to 85% or more of all the markings, even of the lowest third of the classes.

There are two respects in particular in which the correlation does not hold in this table, namely, management and attention to detail of school business as teachers. Several causes may be mentioned which will explain the lack of correlation in this particular in this table. First, these markings represent only one and one-half mark per teacher. This means that these marks cover only the first term of teaching for one half and a second term for the other half of the group. Second, no instruction is given in the management of children and attention to details of school business in the College. Third, the young teacher, facing the numerous situations, naturally emphasizes those things she can do best and that is the instruction. Fourth, management and attention to details are more or less matters of routine that the teacher must learn from experience and which she would naturally find irksome until they are routinized. They do not become routinized in the first term. Fifth, and finally, the correlation does hold in later teaching as indicated by data which are not yet in shape for presentation.

The degree of correlation indicated in this study seems to lead to the unmistakable conclusion that scholarship, measured by the standards of the schools, will indicate success in teaching in after life in spite of the numerous exceptions to the rule that lead so often to the erroneous statement that there is no correlation between scholarship and teaching ability. Whether this same degree of correlation prevails in other vocations, we make no attempt to say.

E. George Payne.

Principal Harris Teachers College,
St. Louis, Missouri.

PROVISION FOR A POST-SENIOR YEAR FOR HIGH SCHOOL SCIENCE TEACHERS

A post-senior year has been provided at the Carnegie Institute of Technology as a part of its professional requirement for those fitting themselves to be teachers of the physical sciences in high schools. About a quarter of the time in this fifth year is devoted to supervised teaching. This practice is given in the institution. A limited number of post-senior students will be paid sufficient for this work to relieve them of the financial burden of the extra year in their course. The plan is somewhat analogous to the hospital service of an interne in the medical profession.

Graduates of other colleges who have taken their degrees in general science and carried some work in psychology or education can usually qualify for the post-

senior year without taking extra courses. At present the outline of the training course provides only for teachers of physics and chemistry, although it is planned later to extend the work to preparation for teaching other sciences. The basis is also laid in education and psychology for advancement to administrative positions since promotion often occurs in this field to a principalship or executive work for those showing the necessary personal traits.

By means of the fifth year of required work the prospective teacher is carried considerably farther in the intensive courses in his particular subject than is possible at colleges which prepare high school teachers with a shorter course. It is believed that this extra training in his science, coupled with the technique of teaching, will equip the science teacher so that he can bring to the high school students the inspiration of the man who knows more than his text. For teaching in college the student is expected to plan for additional graduate years.

Since a large portion of the five-year science teachers' course parallels the four-year preparation for the scientific research option, the graduate is also prepared for positions in government or industrial research laboratories. This research preparation is planned to afford better training for many positions than the regular engineering or liberal arts courses. All general science students study French and German and continue their major science with special instruction in theoretical subjects and laboratory practice to enable them at the time of graduation to undertake immediately many of the problems assigned to physicists or chemists, thus laying a substantial foundation for post-graduate work.

In the bulletin announcing these courses the student is cautioned as to the personal qualifications he should possess. This official description for the vocational guidance of prospective college students is so unusual that it will bear quotation.

"To succeed in the Science Course and in Research, a student should possess unusual ability in Mathematics, Physics, and Chemistry. He must have an interest not only in the practical applications of his major subject, but also in theoretical study and investigation, and should be a student by inclination. He must be patient and painstakingly thorough. If he is also original he should be very successful.

"A student planning to take the teachers' option of this course should, in addition to the above qualifications, possess an attractive personality, teaching ability, and the faculty of clear and forceful expression. He will be more successful as a teacher if he possesses habits of openmindedness, geniality, fairness, and the ability to put himself at the other person's point of view. He should be interested most of all in young men; and to him enthusiasm for his subject and for teaching that subject should outweigh the greater financial attractions of other professions."

J. B. Miner.

Carnegie Institute of Technology,
Pittsburgh.

ABSTRACTS AND REVIEWS

DARWIN OLIVER LYON. *Memory and the Learning Process.* Baltimore: Warwick and York, Inc., 1917. Pp. 184. $2.00.

In this work the author discusses the memory problem from its various aspects —giving, however, special attention to its relation to economy in learning, the best form of imagery to be employed for any special task or procedure, and the uselessness of the various "systems" and mnemonic "devices" still so prevalent in our magazine advertisements.

The structure of Dr. Lyon's book, as he himself points out, is somewhat composite—two of the chapters being essentially experimental in character, while the others either deal with conclusions drawn from the data in question or are wholly theoretical in character. The outcome, however, is that he presents a rather complete discussion of the facts of memory and their practical application.

Chapter 1 opens with a discussion of the divisibility of memory—from the view point of anatomy, physiology, and psychology. He then takes up memory in its relation to association and reasoning. This leads him to a discussion of the association of words versus the association of ideas—or, as the author puts it "literal" learning versus "logical" learning. The chapter closes with a discussion of professional and occupational memories—showing the effects of education, environment, sex, race, etc. Following this the author again takes up imagery and endeavors to show that there are as many ways of comprehending as there are forms of imagery.

The main purpose of Chapter II would seem to be to show just *how we learn—how we memorize.* To do this the author takes up one at a time the four sub-factors of memory, retention, reproduction, representation, and identification.

Chapter III is divided into two parts (1) The Distribution of Time in Relation to Economy in Learning and Retention; and (2) The Relation of Length of Material to Time Taken for Learning. Concerning the first of these, it is shown that in estimating *economy,* not only must we consider the *time spent,* but the degrees of *retention* as well. Concerning the second, tables are presented to show that the relation depends almost wholly upon the division of the time in learning, *i. e.* the distribution of the time intervals. This chapter is an exceptionally definite and valuable contribution to educational processes.

Chapter IV is devoted to an extensive inquiry of several memory problems, such, for example, as the relation of quickness of learning to retentiveness. One of the best points brought out by the author is the demonstration of how complex and elaborate any one problem may be, and what a large amount of work must be performed before any approximately correct answer may be given.

The popular impression among the laity is that the slow but steady worker, even though dull, remembers his work better and longer than the more brilliant student, —a corrollary of which is that those who learn the quickest forget the quickest. However, in so far as reliable statistics have been gathered it has been found that in general the most rapid learner is also the best retainer. A complete expression of the various results obtained with the various methods and materials is obviously impossible in this brief summary. Generally speaking, we may say that those who learn quickly remember longest if the material memorized is "meaningful" or "logical" but that they forget quickly if the material is such as involves the memorizing of motor associations, as is generally the case with digits, words, and nonsense syllables.

As to economy in learning the author shows that before this problem can be intelligently discussed one must first define exactly what he means by "economy." In short, Dr. Lyon shows that to commit to memory is nothing but to make the acquired idea as capable as possible of reproduction; not only that it may be easily and faithfully reproduced but that its relationships may be many and varied, *i. e.* that it may be reached from many sides and through different channels and ideas. The author points out that in every day life not only a one-sided but a many-sided power of recall is demanded. Hence it follows that if a certain method of memorizing results in a more-sided power of recall, this method will deserve more credit than one that secures merely a one-sided reproduction.

The author points out that no determination of economy in any one method can have universal validity, for the reason that the purpose that the learner may have in view differs in various cases. For example, one method may lead rapidly to an errorless first recitation but the permanence of the retention may be low; another method, on the other hand, may lead much less rapidly to a first errorless recitation, but the retention secured by it may have a high degree of "permanicity." In short, economy of learning depends upon the purpose or result desired by the individual doing the learning in question. Hence if any one included in his meaning of memory some certain method, *i. e.*, emphasis on the state of health and mental condition of the individual at the time of his memorizing, the various other methods would for the purpose of discussion remain unused.

Concerning the educational value of psychological research, it is impossible in this space to give a connected and complete account of Dr. Lyon's opinion. As a piece of literature alone this chapter ranks high both because of the felicity of the style and the matter-of-fact and practical argument presented. In some respects it represents the most important contribution of the volume. At first sight it may impress the reader as a criticism—and a rather harsh one—of experimental psychology in general, but closer analysis will show that as a matter of fact the experimental psychologist is here defended and praised. There is also an interesting discussion on mnemonics, showing that on the whole they are of little value.

Taken as a whole this chapter presents us with the most comprehensive discussion of the relation that experimental psychology bears to education that has appeared recently. It will be seen that Dr. Lyon, though thoroughly trained in the Herbartian psychology and pedagogy, does not regard them as the consummate formulation of educational theory. He considers it necessary to supplement Herbart by a more practical knowledge of experimental psychology in its bearings on education. He seems to think that much of the work performed in the psychological laboratories is time wasted, yet he tells us that the educator can not get along without the psychologist, and it is evident that he desired to win for exact psychology a greater amount of attention and interest from pedagogy. It may appear that the author is somewhat too harsh in his criticism of the "results obtained by the experimental psychologist," but there remains ground for abundant praise of the chapter in question, as well as of the work as a whole.

The book is fairly well printed and carries an index of authors as well as a carefully prepared subject index. There are 30 tables (including those in the Appendix) and in the pocket in the front cover are nine large plates in two colors. Many of these plates would be in more convenient form if they had been reduced in size and included in the regular binding of the book.

J. V. Breitwieser.

Colorado College.

HAROLD O. RUGG. *Statistical Methods Applied to Education.* Riverside Textbooks in Education. Boston: Houghton Mifflin Co., 1917. Pp. xviii, 410. $2.00.

What Horace Mann did in 1843 for educational statistics Rugg now has done in 1917. Over half a century ago the need for a quantitative expression of educational conditions made itself felt. But it was not without a great deal of general opposition that Mann was able to accomplish what he did. At the present time, however, we find school people, with few rare exceptions, more than favorable to a statistical treatment of educational facts. Indeed that administrator who has no need for refined methods of evaluating school results is as difficult to find as the now famous needle in the haystack. The book fills a long felt need, and will undoubtedly help trace a new high-water mark in education.

The plan and purpose of the book is outlined in the preface. Rugg has assumed "that it is necessary to equip school men, generally, with a thorough-going knowledge of statistical methods; that in order for them to be discriminating in the use of the various methods in improving their school practice, this large background of knowledge must be developed; and that it is possible to explain rather completely the reasons for and the significance of the principal statistical devices without expressing the explanation in technical mathematical language." When the student turns the last page of the book he is convinced that Rugg has made these statements good.

The first three chapters of the book are devoted to a discussion of the statistical methods in education, the collection of educational facts, and the tabulation of educational data. A brief survey is made of recent quantitative educational literature and its importance is emphasized in the light of administrative and pedagogical-experimental problems. The necessary steps in educational research are then outlined, and the sources for the collection of data indicated. These sources have now been amply supplemented by the author in the current numbers of "The School Review" and "The Elementary School Journal." It is also gratifying to find a very critical presentation of the importance of the questionnaire, and the guiding principles which should underlie its preparation. Questions relating to various methods of tabulation are discussed, and a brief description of the Hollerith system is given. One regrets finding no mention of the excellent Findex system, nor of some of the other labor saving devices, such as the slide-rule, comptometer, Brunsviga, space probably being prohibitive.

The next three chapters initiate one into the statistical treatment of educational data. The frequency distribution, the method of averages and the measures of variability are discussed. The terms "unit," "scale," "frequency," "class interval," etc., are explained and illustrated. The chapter on averages is a very valuable one, especially because of the emphasis on uniform terminology and the discussion concerning the relative merits and demerits of the various averages. The reviewer has not been convinced however, that to calculate the median, $\frac{N}{2}$ should be used in preference to $\frac{N+1}{2}$. In view of Meitzen's discussion this change might lead to erroneous results when interpolating (by formula) the values for the different cases in a given class interval. Two of us are now at work on this problem and our results will be published shortly. In stating the formula for the harmonic mean Rugg gives the one most commonly used,

$$\frac{1}{H} = \frac{1}{N} \Sigma \left(\frac{1}{m}\right)$$

The reviewer has simplified this formula:

$$H = \frac{N}{\Sigma \left(\frac{1}{m}\right)}$$

thus facilitating its understanding and computation. The chapter on the measure of variability is an excellent illustration of how an apparently difficult subject matter may be clearly explained if only the author take sufficient pains to think and write plainly and succinctly. The average deviation, the standard deviation, the quartile deviation, the probable error, the semi-interquartile range, the various coefficients of variation and of skewness are interpreted in the light of the normal frequency surface and their calculation illustrated. It seems to the reviewer that it might have been well to indicate that for few cases the standard deviation should be calculated from the formula:

$$\sigma = \frac{\sqrt{\Sigma f d^2}}{N - 1}$$

according to the method of least squares, and under coefficients of variability a consideration of the formula:

$$C. V. = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

would have added to the value of the paragraphs dealing with those specific items.

In the next two chapters the normal frequency surface and the use of the normal frequency curve in education are considered. In using the book with students one misses a discussion of the ogive and its relation to the normal frequency surface. Especially is this true in connection with a study of percentile grading. And here, if I may be permitted to digress, a good portion of the recent work with percentiles is invalid because the ogive and not the frequency surface has been utilized as the base of reference. The major portion of the chapter is devoted to probability, and every bit of it is worth while, since probability is the foundation upon which our whole statistical structure is based. As homely as the comparison may appear, it may be said with justice, that every statistical problem in education brings us back to the question of "tossing pennies." The chapter on the use of the normal frequency curve in education is replete with applications to everyday school problems.

Chapter IX is devoted to a consideration of the measurement of relationship, or correlation. The various methods are indicated and practical applications made. Galton's graphic method, the Pearson coefficient by the long and short methods, the Spearman foot-rule and rank differences formulae, the Pearson *eta* method, the four-fold table, unlike signs and contingency methods are all given. No mention is made of Yule's recent improvement of his Q method. His present coefficient of colligation (ω) is superior to some of the other "association" methods, and far easier to compute. The regression coefficient and the regression equation, their meaning and significance are explained. In the discussion of non-linear regresssion only one of the *etas* is mentioned, η_y, nothing being said regarding the other *eta*, η_x, which is equally important. Thus in one of W. Brown's articles (Brit. J. Psychol. Vol. 6,

p. 236–237), the Pearson coefficient is $+.033\pm.030$, *eta y* is $+.323\pm.027$, and *eta x* is $+.182\pm.029$. This fact should be taken into account when using Blakeman's test for linearity. No space is given to Spearman's correction formula, Rugg being skeptical as to its value. He would be inclined to agree with those who maintain that by making measurements more efficient, correction will be unnecessary. But how can we tell whether measurements are really efficient unless we use a correction formula? To quote Spearman, "The suggestion is like telling a man to brush his coat until it is clean but never look whether it is so." The corrected coefficient is especially necessary when attempting to find the true probable error of a correlation coefficient. True P. E. $= \frac{\text{P. E. r} \times r_{pq}}{r}$. These remarks, however, are not meant to detract from Rugg's very masterly presentation of the measurement of relationship, for he has crowded within a comparatively few pages material which is practically inacessible to the average school man.

The last chapter of the book is in many respects its most unique feature. Everyone interested in "getting facts across" should not fail to read these valuable pages. Figures mean nothing to the statistically uninitiated unless presented in graphic form. These fifty pages are a gold mine of illustrative material.

Other features of great value are a tabular review of quantitative studies on school administration, together with a bibliography of 165 titles, brought to date; a selected bibliography on statistical methods; a set of illustrative problems at the end of most of the chapters; a summary of formulae and symbols used in the text; and ten tables to facilitate computation. There are several typographical errors in the text and tables which a revised edition will eliminate.

The editor of the Riverside Textbooks in Education has not in the least overdrawn the merits of the book when he stated that "The volume represents a very successful attempt to produce a book which will apply the mathematical theory of statistical work to educational problems, and as such it should find a hearty welcome from the teachers of education in universities, colleges, and normal schools, educational investigators generally and school officials interested in making the best of statistical data and displaying the results to their supporting public in the most effective graphic form."

SAMUEL C. KOHS.

Stanford University.

The Journal of Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD.

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

NOTICE TO SUBSCRIBERS. Notify the publishers, Warwick & York, Inc., of any change of address; as copies lost in transit because of subscriber's removal to another address will not be duplicated.

THE JOURNAL is issuing a series of Monograph Supplements for the publicaiion of more extensive researches in educational psychology. For details prospective contributors may apply to H. E. Buchholz, 10 East Center St., Baltimore, Md.

The price of the JOURNAL is $3.00 per year of ten monthly issues, if paid in advance. Single issues are 50 cents each. Back volumes are $3.50 in original wrappers and $4.75 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME IX APRIL, 1918 NUMBER 4

EDITORIAL

Science and Sentimentality

In this day when nations are pouring out their life blood on the battle-field without stint, when we ourselves are sending millions to the front, when our friends and relatives are suffering from wounds and disease, it would seem that every sane individual would welcome and eagerly further any investigations of scientific men to better understand and cope with the problems of healing. Yet, such is the perversity of human nature, a certain class of intellectual people are doing their utmost to hinder scientific experimentation to facilitate the treatment of wounds and disease. This malevolent influence has recently shown itself in the endeavor to block the American Red Cross in its appropriation of funds for "experimentation upon living animals for the purpose of finding methods of prevention and remedies for new and strange diseases among soldiers." "This money," says the general manager of the Red Cross, "was used for the direct and immediate purpose of finding ways to prevent or cure wounds and sickness among American soldiers.' It develops, however, that there are large numbers of earnest Red Cross members who have

sincere convictions against the use of animals for discovery of remedies for sickness. We recognize that it should be an obligation of the Red Cross management to show deference to such honest conviction." Fortunately a private individual offered to supply the necessary funds for carrying on this work, and to reimburse the Red Cross for the money already spent. "The War Council," continues the manager, "decided to accept this offer without in any way taking a position either for or against the question in controversy, but because they do not wish their acts to be considered in conflict with the sincere convictions of Red Cross members."

It is with great pleasure that we quote the following paragraphs from an editorial article in the *New York Times.*

"It will be with regret deeply tinged with indignation that all sane and reasonably enlightened people will hear of the decision by the heads of the Red Cross not to use for animal experimentation—often and almost always incorrectly called 'vivisection'—the money hitherto appropriated by them for that purpose. The decision may mean avoiding the loss of a few contributions to the Red Cross funds, but it also means the triumph of vicious ignorance over common sense, and it will encourage to further efforts the members of the most detestable and not the least dangerous group of men and women to be found in the United States.

"And the Red Cross cautiously says that it does not take sides for or against 'vivisection!' Such caution is reprehensible—is utterly unworthy of that great and beneficent organization. It should take sides, standing for right and against wrong. The immediate profit of doing anything else or less will be dearly bought in future loss of both money and respect. Red Cross money, in the amount that was proposed, could not possibly have been better invested than in the establishment of a biological laboratory near the scene of war for the study of the maladies of soldiers which this sort of research has not yet conquered. It was weak, and worse than weak, for the Red Cross to heed the hysterical shrieks and the monstrous charges of venality and murder that came from a few people whom it strains charity to call deluded or insane."

It would seem that the chief instruments in this obscurantist campaign are the *Christian Science Monitor* and certain individuals prominent in that organization. From people who deliberately disregard the facts of experience, and who express a profound contempt for medical science and natural science of all kinds, nothing more is to be expected. But the sinister aspect of the whole affair is the influence that these people are able to exert on such a respresentative organization as the Red Cross. This is only possible because of a widely spread mushy sentimentalism for animals, which has been fostered by many agencies, one of which is the school. Many times as much effort is devoted in the schools to arousing in pupils the emotion of sympathy for animals as to awakening a whole-hearted appreciation of and enthusiasm for the beneficent effects of biological and sanitary science. Not that pupils should be less humane but that a better and more effective type of humaneness can be attained through an understanding of the benefits of science. In the present crisis it is the duty of every scientific man to uphold the standards of scientific investigation with all the zeal of a religion, and to see to it that his influence is used to the utmost to stamp out any such pernicious and unpatriotic sentimentalism as that which brought pressure to bear on the Red Cross. J. C. B.

NOTES AND NEWS

The matter of the reorganization of education in England moves slowly forward, but not entirely to the satisfaction of those who are zealous for a greater amount of attention to science. Sir Ray Lankester has recently published an article commenting on the report of the government committee regarding the position of science in secondary schools and universities. He welcomes the proposal of the committee to secure equality of opportunity to all branches of learning, but objects to turning the administration of education back to the schools and universities as now organized. "The Committee," he says, "instead of rescuing education from the professional vested interests of the classical schoolmasters, hands back the victim, after many professions of good-will, to the tender mercies of those who are banded together to starve, torture and discredit her, and remorselessly to maintain the domination and the pecuniary allurements of the 'classical system.' "

The American Association of Clinical Psychologists was organized at Pittsburgh, on December 28, 1917. The membership includes men and women holding the doctorate in psychology, who are engaged in the clinical practice of psychology in the United States. The forty-five charter members are chiefly directors of clinics, of bureaus of child welfare, of institutional laboratories; engaged in army service, as mental examiners of recruits; or connected with courts, hospitals and schools. The objects of the Association are to promote an *esprit de corps* among psychologists who have entered the practical field, to provide media for the communication of ideas, to aid in establishing definite standards of professional fitness for the practice of psychology, and to encourage research in problems relating to mental hygiene and corrective education.

The Council of the Southern Society for Philosophy and Psychology has decided, on account of the general situation and of the number of members of the Society who are engaged in various forms of national service, to abandon the annual meeting scheduled to be held at Peabody College, Nashville, this spring.

At the meeting of the New York Branch of the American Psychological Association, April 22, 1918, the following papers were presented: "Mental Effect of Outdoor versus Indoor School Life," by W. A. McCall; "Vocational Tests for Retail Saleswomen," by Elsie Oschrin; "Study of Three Cases of Special Disability in Reading," by Sarah Fisk; "Transfer of Habits in the White Rat," by H. A. Ruger; "Notes with Illustrations Showing the Development and Application of a Child's Number Interest," by the Child's Mother.

The Carnegie Institute of Technology announces ten graduate fellowships and assistantships in salesmanship research and in lines

of non-selling work in department stores. These positions not only yield a handsome stipend to the holders, but carry with them unequalled facilities for study in vocational guidance. Furthermore, in order to encourage young men and women to enter the filed of vocational education as teachers and supervisors about twenty $500 scholarships will be provided for the most capable seniors and juniors who present themselves for this work.

Announcement is now made that Dr. Guy Montrose Whipple has resigned from the University of Illinois in order to accept permanent appointment to the faculty of the Carnegie Institute of Technology as professor of applied psychology and director of a new bureau of educational research. In this capacity as director of educational research, Dr. Whipple will carry forward scientific studies in engineering and technical education, including problems of curriculum, methods of instruction, selection of students, and the like. Dr. Whipple will continue to serve as acting director of the bureau of Salesmanship Research as long as Dr. Scott continues to be engaged in war work.

Dr. James Burt Miner is acting as head of the Division of Applied Psychology at the Carnegie Institute of Technology during the absence of Dr. W. V. Bingham on war work in Washington. Announcement is made of Dr. Miner's promotion to the rank of associate professor.

Dr. L. L. Thurstone has been advanced to the rank of assistant professor at the Carnegie Institute of Technology.

Dr. A. J. Beatty, assistant to the director of the Carnegie Bureau of Salesmanship Research, will on June 1st become director of education of the American Rolling Mills Company at Marietta, Ohio.

Dr. Kate Gordon has been granted leave of absence from the Carnegie Institute of Technology for the fall quarter to enable her to carry out for the California State Board of Control a psychological investigation of the children in certain of the state institutions.

Dr. Beardsley Ruml has been given leave of absence from the Carnegie Institute of Technology to devote his full time to the direction of the work of the Trade Test Standardization Division of the Committee on Classification of Personnel in the Army. Dr. L. L. Thurstone has been granted half time leave for similar work. Dr. T. J. Kirby has been granted half time leave from the University of Pittsburgh and is working with Dr. Thurstone, Mr. L. C. Toops of the University of Ohio, and Dr. J. Crosby Chapman who is in charge of the Pittsburgh station of this Trade Test Standardization Committee. The purpose of the development of these standardized trade tests is not to discover which trade or occupation a soldier should be trained to follow. It is rather to measure the degree of

trade skill which he already has. The question is not one of "guidance" but of assignment of men to those duties of a technical sort which their civilian occupations have equipped them to follow to advantage in the army. Oral and performance tests of carpenters, pattern makers, vulcanizers, automobile engine repairmen, truck drivers, electricians, etc., have been developed, standardized and introduced into army procedure. Tests for skill in more than a hundred other trades of importance in a modern army remain to be developed and standardized. About twenty mechanical engineers, civil service experts, employment managers and psychologists are engaged in the preparation of these trade tests, working under the immediate supervision of Dr. Ruml at Newark, New Jersey, and under the more general direction of Dr. Bingham who is executive secretary of the Committee on Classification of Personnel in the Army, with headquarters in the office of the Adjutant General at Washington.

Professor E. A. Kirkpatrick, of Fitchburg, Mass., announces that he is now able to supply to psychologists and educators pictures of Professor John Dewey. He has been able to furnish pictures of other leaders in education for some time, but has had to disappoint the many inquirers for Professor Dewey's picture.

Mr. Murray Dalman, assistant principal at Emmerich Manual Training High School, has been appointed director of a department of educational reference and research recently created at Indianapolis.—*School and Society.*

L. O. Smith, a graduate student in education and psychology at the University of Iowa, has been selected as director of educational research in the Omaha schools.

Dr. Buford Jennette Johnson, Ph.D. (Hopkins '16), has resigned her position as assistant psychologist in the Laboratory of Social Hygiene, Bedford Hills, N. Y., and has accepted an appointment as research assistant in the Bureau of Educational Experiments, New York City.—*Science.*

Harold Ernest Burtt has been appointed instructor in psychology at Harvard University.

PUBLICATIONS RECEIVED

THOMAS ALEXANDER. *The Prussian Elementary Schools.* New York: The Macmillan Company, 1918. Pp. viii, 571. $2.50.

It is desirable to understand one's enemy, and no one can understand the German people who fails to take account of the German schools. Much has been written of the German higher schools, but these are intended only to increase the efficiency of the upper classes. The Germans, more than any other people in history, have used their elementary schools to develop a unified and solidified public sentiment for the support of a definite type of government and a specific national ideal. The aims of German elementary education and the detailed procedure by which these aims have been realized are admirably set forth in the present valuable book. "The elementary schools of Prussia have been fashioned so as to make spiritual and intellectual slaves of the lower classes . . . . . The Prussian elementary school is the best in the world from the point of view of the upper classes of Germany. From the point of view of the lower classes it is the worst system, for it takes from them all hope of improving their condition in life." The first twelve chapters of the book give a clear and accurate account of the more general features of Prussian schools, such as their development, administration, school attendance, school management, preparation, salaries and pensions of teachers, and the organization and courses of study. The remaining sixteen chapters deal with general and special methods of instruction and the organization of subject matter in religion, mother tongue, mathematics, history, geography, natural sciences, domestic economy, singing and drawing, and manual and physical training. The book will unquestionably take its place as the standard authority on the subject.

J. MACE ANDRESS. *The Teaching of Hygiene in the Grades.* Boston: Houghton Mifflin Company, 1918. Pp. xii, 177. $0.75.

Sane and normal living is one of the most important ideals of education, yet it is only in recent years that much attention has been paid to hygiene in the elementary schools. This little volume gives an excellent statement of the arguments for the emphasis of hygiene in elementary education, shows how far from satisfactory is the present status of the subject, indicates goals of instruction, makes valuable suggestions on methods of teaching, and discusses several important problems in hygiene. Among the latter are alcohol, tobacco, colds, tuberculosis, cleanliness, exercise, care of the teeth, proper foods, pure milk, elimination of waste, care of the eyes, care of babies, sex hygiene, the house-fly, the mosquito, the cat, rats and mice, and the prevention of fires and accidents. There are references to standard books under each topic.

FRANK W. BALLOU. *Arithmetic. The Courtis Standard Tests in Boston, 1912–1915. An Appraisal.* Boston Public Schools, Bulletin No. 10 of the Department of Educational Investigation and Measurement, 1916. Pp. 48. Seven cents.

In those schools where standard tests have been given regularly the pupils show as much as seventeen per cent. superiority over pupils who were tested for the first time. The pamphlet gives valuable norms and suggestions for procedure.

JOHN M. BREWER. *The Vocational-Guidance Movement: Its Problems and Possibilities.* New York: The Macmillan Company, 1918. Pp. xi, 333. $1.25.

This is a thoughtful and scholarly consideration of a very important subject. The author surveys some of the problems of vocational guidance, traces the history of the movement, shows its relation to educational guidance, points out many of the pitfalls into which advocates of guidance have fallen, and outlines a program for its development. He is extremely skeptical regarding the value of psychological tests in vocational guidance, and subjects to very shrewd criticism the proposals of Muensterberg, Lough, Scott and others to determine capacities by responses to artificial situations. This effort he thinks is the result of looking upon capacities as fixed entities, which is a survival of the old "type" or "faculty" conceptions in psychology. All that vocational guidance can do is to secure a rational type of school, to bring teachers into sympathetic contact with the inner lives of their pupils, to co-operate with the child in choosing and starting in a life career, to endeavor to improve the conditions of employment, to stimulate the child's interest in his career and help him see its opportunities, and to give him such training in his chosen occupation that he may work to his own best advantage. While the author may be unduly pessimistic about the outcome of psychological and educational tests, his strictures deserve careful consideration, and the general tone of his book is sound and salutary.

CATHERINE C. CONWAY. *Performance Norms for Thirteen Tests.* New York State Board of Charities. Eugenics and Social Welfare Bulletin, No. viii, 1917. Pp. 142.

The tests studied in this monograph are the Knox cube, number cancellation, recall of objects, grouping of objects, peg design, story reproduction, syllogisms, detail drawings, balance nickel, peg board, tower, and boat. The total number of children examined was 757.

JOHN EDGAR COOVER. *Experiments in Psychological Research at Leland Stanford University.* Leland Stanford Junior University Publications, Psychical Research Monograph No. 1, 1917. Pp. xxiii, 641. $3.50.

This stately volume represents the first fruits of the Stanford $50,000 fund for the investigation of spiritualism and allied psychic phenomena. The attitude assumed by the investigator is one of strict scientific inquiry, and an attempt has been made to subject certain occult phenomena to experimental scrutiny. The first topic, thought-transference, is discussed from a theoretical point of view, and experiments on guessing of numbers, guessing of playing cards, and the "feeling of being stared at" are reported. Other topics considered are experiments on subliminal impression, mental habit, experiments in sound assimilation, and investigations of pseudo-prophecy, automatic writing, and long-distance thought-transference. The conclusions negative any mystical assumptions, but bring out some interesting phases of human mental processes.

GEORGE S. COUNTS. *Arithmetic Tests and Studies in the Psychology of Arithmetic.* Chicago: University of Chicago Press, 1917. Pp. iv, 128. $0.75.

This monograph presents a detailed study of the results of the surveys of the Cleveland and the Grand Rapids schools, so far as the arithmetic was concerned. Fifteen tests, constructed on the basis of the Courtis Tests, but involving a greater

variety of situations, were given to 834 classes in grades III to VIII. The data are organized under general results, types of errors, age groups, promotion groups, and race groups. The types of errors afford interesting material for the teacher.

ISOBEL DAVIDSON. *Real Stories from Baltimore County History.* Baltimore: Warwick and York, 1917. Pp. 282. $1.00.

Local history, as well as local geography, is now advocated for the grades, and the present book is the result of a remarkably successful effort to form a collection of local history stories by pupils and teachers. The stories fall into three groups, the early pioneer days, the colonial days, and the time from the stage-coach days to the present. Teachers will find this book interesting and suggestive not only for the material it contains but also as a model for similar local collections.

KARY C. DAVIS. *School and Home Gardening.* Philadelphia: J. B. Lippincott Company, 1918. Pp. xvii, 353. $1.28.

This splendidly illustrated manual will be of great value to those who are directing or planting school or home gardens. There are chapters on planning the garden, tools, soil, germination, drainage, vegetable and flower gardens, trees, and vines, a garden calendar for northern and for southern states, and suggestions for club work and garden contests.

JAMES FAIRGRIEVE. *Geography and World Power.* New York: E. P. Dutton and Company, 1917. Pp. viii, 356. $1.50.

Man boasts that he has triumphed over nature, yet it is nature that makes man's very existence tolerable. Throughout the world's history national power has been gained only through utilizing the forces of nature. The author shows that from the time of the Egyptians to the present day national power has depended on geographical factors, and it is incumbent upon us to realize what the favorable factors are and to make the best possible use of them. It is a most interesting book.

FREDERIC HENRY GERRISH. *Sex-Hygiene: A Talk to College Boys.* Boston: Richard G. Badger, 1917. Pp. 51. Sixty cents.

This is a lecture on sexual continuence that has for many years been delivered annually before the freshman class of Bowdoin College. The dangers of venereal disease are portrayed, and ideals of purity and self-control are urged.

JOHN M. GREGORY. *The Seven Laws of Teaching.* Boston: The Pilgrim Press, 1917. Pp. xii, 129. Seventy-five cents.

This little book by the founder of the University of Illinois is presented in revised form by Dr. W. C. Bagley and Warren K. Layton, of the School of Education of that institution. It sets forth with luminous clarity and vigorous directness those essentials of teaching which have stood the test of the most rigorous psychological criticism. It would be difficult to find so much sound wisdom packed away in such small compass.

HERBERT S. JENNINGS, JOHN B. WATSON, ADOLF MEYER AND WILLIAM I. THOMAS. *Suggestions of Modern Science Concerning Education.* New York: The Macmillan Company, 1917. Pp. vii, 211. $1.00.

The four addresses here presented give us the reactions of the specialist in biology, psychology, psychopathology and sociology to fundamental questions in education.

Professor Jennings makes a serious arraignment of present school practices from the point of view of the welfare of the race, and sounds a call for greater attention to and study of school conditions. Professor Watson tells of his experimental studies of infants, and pleads for a closer connection between the laboratory and the school. Professor Meyer urges a greater individualization of the instruction to prevent the nervous breakdowns which are all too frequent, and Professor Thomas emphasizes the need for a science of social behavior to furnish a basis for educational proeedure. The book is tremendously stimulating, and deserves wide and thoughtful study.

PERCY GAMBLE KAMMERER. *The Unmarried Mother: A Study of Five Hundred Cases.* Boston: Little, Brown and Company, 1918. Pp. xv, 342. $3.00.

Self-control in sex relationships is generally considered one of the most important topics in moral education. Latterly society is beginning to inquire as to its responsibility for the upbringing of the illegitimate child. The conditions brought about by the war have intensified interest in both of these topics. In the present book we have a detailed, first-hand study of the conditions which lead women to become unmarried mothers. The author recommends laws for the protection of mentally abnormal women; laws for the better care of the child; improvement of recreational facilities; and a more rational teaching in regard to sex hygiene.

STERLING ANDRUS LEONARD. *English Composition as a Social Problem.* Boston: Houghton, Mifflin Company, 1917. Pp. xiii, 202.

The principal suggestions in this book in regard to teaching English composition are the idea of the social character of the composition, such handling of organization problems that children's minds will work best in mastering them, and the distinction between matters to be fixed as habits and those to be developed through conscious reflection. The whole aim is to ensure expressional skill in children, and the author's methods will command the unqualified approval of the psychologist.

RICHARD SWAN LULL. *Organic Evolution: A Text-book.* New York: The Macmillan Company, 1917. Pp. xviii, 729. $3.00.

This is a magnificently complete, thorough, authoritative account of the evidence for the evolution of animal life drawn especially from fossil remains. The book is the outgrowth of a series of lectures which has been given annually to students of Yale University for more than ten years. In it the author discusses the history of the evolutionary idea, the distribution of animal life, the various theories of the mechanism of evolution, and the evidences of evolution derived from the life cycle of animals, from adaptations to their environment, and from the records left by organisms in the various geological strata. It is a useful and comprehensive handbook for all questions connected with evolution.

STEWART MCDOWELL. *Seven Doubts of a Biologist.* New York: Longmans, Green and Company, 1917. Forty cents.

In this little book the author, who represents himself as a trained biologist, takes up one after another the religious doubts that are supposed to be raised by the study of modern biology. It is to be feared that the author merely knocks over a man of straw, for either the doubts are not fairly stated, or they are answered by such inconclusive arguments as would have little weight with any open-minded doubter.

C. R. MAXWELL. *The Observation of Teaching.* Boston: Houghton Mifflin Company, Riverside Educational Monographs, 1917. Pp. xi, 120. $0.75.

This monograph is designed as an aid for students in training and for all who are interested in analyzing the teaching process. To get adequate value from observation of teaching one must have a clear idea of what to look for. To this end the author outlines the desirable characteristics of the teacher, the appearance and attitude of the pupils, the procedure in various types of lessons, and the physical features of the schoolroom. The practical difficulty with this scheme is that it is too bulky to be readily applied.

Mental Examinations. Eugenics and Social Welfare Bulletin, No. 11, New York State Board of Charities, 1917. Pp. 73.

There is an account of the examination of 2,142 orphan asylum children, 607 delinquent girls, 183 reformatory women and certain classes of school children by the Binet tests, and the re-examination of 37 children ten months after the first examination.

IRVING ELGAR MILLER. *Education for the Needs of Life. A Textbook in the Principles of Education.* New York: The Macmillan Company, 1917. Pp vii, 353. $1.25.

This book is the outgrowth of the author's "Psychology of Thinking." The biological point of view is emphasized throughout, and the attempt is made to interpret the meaning and aim of education in terms of individual and social values, to trace the essential steps in the development of the child, to apply the ideals of individual and social needs to the construction of the curriculum, to formulate principles of method on the basis of the child's self-activity, and to point out the functions and qualifications of the teacher. The book is not profound, but it is well-conceived, and takes account of the best and most recent educational literature.

WALTER SCOTT MONROE. *Development of Arithmetic as a School Subject.* Washington: Bureau of Education, Bulletin, 1917, No. 10. Pp. 170. Twenty cents.

This is a careful and authoritative study of the history of the teaching of arithmetic in American schools. The earliest accounts of arithmetic in the colonies are recorded, the texts of Dilworth, Pike and Adams are analyzed, the cyphering methods are described, Warren Colburn and the Pestalozzian influence are discussed, and the host of more recent arithmetics from Emerson, Davies and Ray to the experimental studies of Rice, Stone and Courtis receive adequate attention. It is an admirable piece of work, and deserves the widest circulation.

WALTER S. MONROE, J. C. DEVOSS AND F. J. KELLY. *A Teacher's Handbook on Educational Measurements, Reading, Writing and Arithmetic.* Kansas State Normal School: Studies by the Bureau of Educational Measurements and Standards, No. 5, 1917. Pp. 66.

This is a preliminary survey of the recent experimental work in the school subjects mentioned, and was intended for local distribution. The work has been expanded and published under the title of "*Educational Tests and Measurements*" by Houghton Mifflin Company.

MARGARET B. OWEN. *The Secret of Typewriting Speed.* Chicago: Forbes and Company, 1917. Pp. 158. $1.00.

The author of this book is said to have won the world's typewriting speed championship four times, and the last time to have written 143 words per minute for an hour. It would be extremely interesting if she could tell how she did it, but she cannot, or at least has not done so here. The book contains much good practical advice to the typewriter, and will undoubtedly be of service to those specializing in this field.

JOHN FRANKLIN REIGART. *The Lancasterian System of Instruction in the Schools of New York City.* New York: Teachers College, Contributons to Education, No. 81, 1916. Pp. v, 105. $1.25.

Perhaps the most valuable portion of this interesting study is the chapter on the method of teaching which Lancaster employed in the instruction in specific school subjects. The chapter on the organization and administration of the monitorial schools, and that on moral and religious training throw much light on the New York Free Schools in the early half of last century.

CHARLES L. ROBBINS. *The School as a Social Institution, An Introduction to the Study of Social Education.* New York: Allyn and Bacon, 1918. Pp. xxv, 470.

It is generally recognized that the social aspects of education need greater emphasis and more consistent attention, but one difficulty heretofore has been the lack of suitable treatises on the subject. The present volume will contribute to removing that lack, and should receive wide-spread consideration. Among the topics treated are the evolution of the idea of the school as a social institution, social ideals and conditions affecting the school, reaction upon society, health protection and the schools, the selection of the socially fit, vocational and educational guidance, the moral significance of the schools, community centers, the social elements in the curriculum, social aspects of method, and the recruiting of teac ers as a social problem. The treatment of these topics is sympathetic and progressive and the book points to the greatly increased usefulness of the schools.

BERTRAND RUSSELL. *Political Ideals.* New York: The Century Company, 1917. Pp. 172. $1.00.

A collection of essays on topics of social import. The first, which gives the title to the book, examines the foundations of government, and points the way to a world democracy. Others are concerned with capitalism, pitfalls in socialism, individual liberty and public control, national independence and internationalism. Throughout the discussion the author seeks the greatest individual liberty with the greatest advantage to the common weal.

CARL E. SEASHORE, Editor. *University of Iowa Studies in Psychology. No. VII.* Psychological Monographs, Vol. 25, No. 2, 1918. Pp. 163.

The most significant paper for the educational psychologist in this collection is that on "Correlation of Factors in Musical Talent and Training" by Carl E. Seashore and George H. Mount. In this the authors bring together materials which have been collected in the Iowa laboratory for the past ten years regarding tests of musical ability, and endeavor to determine statistically the relationship between the various tests that have been employed. This is a most important discussion of the

fundamentals of musical testing, and can be cordially commended to the attention of all those who are interested in tests of any sort. The highest correlation (.61) was between singing interval and singing keynote. The highest correlation between pitch discrimination and fourteen tests was that with tonal imagery (.52).

HENRY LESTER SMITH. *A Survey of a Public School System.* New York: Teachers College, Contributions to Education, No. 82, 1917. Pp. xii, 304.

A detailed account of a school survey of Bloomington, Indiana. In the chapter on the achievements of pupils, to which over one hundred pages are devoted, we have scores on the Stone, Courtis, Haggerty and other tests in arithmetic; handwriting measured by both the Thorndike and the Ayres scales; the Rice, Courtis, and Buckingham tests in spelling; the Courtis Standard Tests in Composition, scored by the Hillegas scale; drawings scored by the Thorndike scale with the Childs extension; and reading scored by the Thorndike scales. In many respects this is the most scientific of school surveys.

FRANK N. SPINDLER. *The Sense of Sight.* New York: Moffat, Yard and Company, 1917. Pp. xv, 156. $1.25.

This is one of the books in the series entitled "Our Senses," edited by Professor George Van Ness Dearborn. It purports to give a popular account of the facts of vision. There is a real need for such a book, but in it one expects a reasonable completeness and accuracy of statement, and in both respects the present volume leaves something to be desired. There is no discussion of the chief visual defects, the figures are misleading, and much emphasis is laid on the generally discredited theory of visual types. It is to be hoped that other volumes of the series will be more scientific.

LINA ROGERS STRUTHERS. *The School Nurse.* New York: G. P. Putnam's Sons, 1917. Pp. xiv, 293. $1.75.

There is a growing recognition of the school as an important agency of social betterment, and one of the most potent faos in this activity is the school nurse. The present volume by the former superintendent of school nurses of New York City will be of value to school nurses, teachers, and especially to parents. There is a history of the school nurse movement, plans for the organization and administration of school nursing, school clinics, children's diseases and their treatment, common physical defects, decayed teeth, and the qualifications of the school nurse. The book is a distinct contribution to an important phase of school service.

E. H. STURTEVANT. *Linguistic Change. An Introduction to the Historical Study of Language.* Chicago: The University of Chicago Press, 1917. Pp. x, 185. $1.00.

This is a delightful and informing little book for those who take pleasure in tracing out peculiarities of language. While scholarly, it is not technical, and can be read with enjoyment by an intelligent high school graduate. There is an introduction on the nature of language, chapters on changes of form, change of meaning, change in vocabulary, change in syntax, language and dialect, and the trend of linguistic development.

ROLLA M. TRYON. *Household Manufacturers in the United States, 1640–1860.* Chiago: The University of Chicago Press, 1917. Pp. xii, 413. $2.00.

The sub-title of this valuable book is "A Study in Industrial History." The

author first considers the factors affecting household manufactures in the early days, such as England's colonial policy and the efforts of the colonies to foster industries, then describes the status of household manufacturing in the seventeenth and eighteenth centuries, considers the changes brought by the revolution, gives detailed statistics of a year's output in 1810, and sketches the transition to shop- and factory-made goods. Only the perusal of such a book brings home the vast changes in living which the past century has witnessed.

GEORGE RANSOM TWISS. *A Textbook in the Principles of Science Teaching.* New York: The Macmillan Company, 1917. Pp. xxvi, 486. $1.40.

This splendid survey of the teaching of science calls forth the unqualified approval of the critical student of education. Rarely does one find such grasp, such breadth of view, such sanity and clarity of presentation as is evidenced here. The significance of science in life and in the curriculum is ably set forth; the doctrine of mental discipline is stated with due caution, and with full appreciation of the findings of modern psychology; the suggestions on method are ample without being burdensome; and the bibliographies are exceptionally complete and well chosen. The chief individual sciences considered are biology, geography, physics, and chemistry. There is an edifying chapter on "general science" courses in which the project method is given the preference over the informational method. A final chapter on examinations and tests emphasizes the variability of teachers' marks according to recent studies, and recommends a ranking method rather than the usual percentage method of rating. The book shows a study and appreciation of the results of recent experimental and statistical investigations of education that is to be commended to other writers on secondary education.

THE WORLD BOOK. *Organized Knowledge in Story and Picture.* Edited by M. V. O'Shea, Ellsworth D. Foster, and George H. Locke, assisted by one hundred and fifty distinguished scientists, educators, artists and leaders of thought in the United States and Canada. Chicago: Hanson-Roach-Fowler Company, 1917. Complete in eight volumes. Vol. 3, Crow-Gloucester, pp. 1649–2512. Vol. 4, Glove-Lemay, pp. 2513–3376. Vol. 5, Lemberg-New Year, pp. 3377–4190. Vol. 6, New York-Rice, pp., 4192–5008. Vol. 7, Richard–Tides, pp., 5009–5808. Vol. 8, Tie–Zwingli, pp., 5809–6528.

The present volumes of this magnificent educational compendium bear out the promise of the earlier ones. It is indeed a treat to pick up these sumptuous volumes and explore the riches that lie at hand for both teacher and pupil, young and old. Is the child interested in nature study? Then he will find here three pages on minerals, five pages on fishes, several pages on birds, reptiles, mammals, insects, plant life, and nine pages on nature study itself. If it is some difficulty in fractions that bothers the pupil he will find a clear explanation under that head. In geography, astronomy, history, literature, and science the high school pupil will find easily and quickly just what he most needs. The elementary teacher will be able to use the work to advantage in the preparation of almost every lesson. On topics of current interest, such as the geography of the present war, no reference work can be found which will compare with this in timeliness and completeness. In short, for any educational or informational purpose whatsoever one could scarcely err in recommending this work too highly.

VOLUME IX MAY, 1918 NUMBER 5

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

AN ABSOLUTE POINT SCALE FOR THE GROUP MEASUREMENTS OF INTELLIGENCE. PART I.*

ARTHUR S. OTIS

Surgeon General's Office, Washington, D. C.

CONTENTS

*The writer is indebted to Dr. Lewis M. Terman of Stanford University for many helpful suggestions during the making of this study.

I. Introduction

Purpose—The purposes of this study are:

(1) To construct a scale for the measurement of general mental ability, such scale being:

(a) suitable primarily for administration to the pupils in grades 4, 5, 6, 7, and 8 of the elementary school,

(b) capable of being administered to *groups* of at least 50,

(c) so constructed that the scoring is both rapid and, as far as possible, free from the error of the personal factor,

(d) built upon the general plan for an Absolute Point Scale outlined by the writer (see Ref. 10). Such construction would involve the "validation" and "graduation" of the tests, the determination of the probable error of a determined measure of general mental ability, etc.

(2) To investigate the correlation of the mental abilities tested by the scale.

It is not deemed feasible, in the space to which this article is limited, to attempt to discuss the nature of intelligence, such as it is presumed to measure by the present scale, nor the various definitions of intelligence which have been given or may be implied by the various 'intelligence scales' in present use. These subjects will be touched upon in various connections in the discussion. The writer will therefore proceed immediately to describe the manner in which the present scale was constructed.

II. The Tests

Requirements of a Scale for Mass Testing.—The chief object of testing in groups, of course, is economy of time. One of the most essential means of accomplishing this purpose is that the responses required be very simple. This makes for speed both in the administration and in the scoring of the tests. It has been the aim, therefore, so to arrange the tests in the present scale that there would be only one correct answer to each item and that this might be indicated merely by making a letter or figure or drawing a line. Where convenient, provision was made for the responses to be placed in a single column in which case the papers may be scored with dispatch

by the use of scoring forms. When every answer is either right or wrong, a large amount of time is saved that might be necessary otherwise to determine the value of partially correct answers. Moreover, under these conditions the tests must be scored in the same way by all investigators, this assuring comparability.

The ideas for the tests have been derived from various sources, chiefly, perhaps, from the Stanford Revision of the Binet Scale. (see Ref. 14). The general character of most of them is no doubt familiar.* In substance the remaining tests were designed especially for this study.

Description of the Tests.—The scale was compiled in duplicate. There were, in other words, two complete tests of each kind. The two tests of each kind were made as nearly alike as possible without using the same material. In each scale the tests were constructed as follows:

The Spelling Test† consisted of fifty pairs of words in two columns. The words of each pair consisted of the correct and incorrect spelling of a single word. In some cases the first spelling was the correct one and in some cases the second was the correct one. The pupil was required to indicate by the letters, F, S, or N, placed in a parenthesis opposite the words, as shown in the sample in Appendix I, whether the first or the second was the correct spelling, nor neither spelling was correct.

The Arithmetic Test consisted of 16 problems in which the computation was made as easy as possible and the emphasis thus placed upon reasoning.

The Synonym and Antonym Test consisted of 50 pairs of words as shown in the sample. The pupils were required to indicate by the letters, S and O, whether the words of a pair meant the same or the opposite.

The Proverb Test consisted of 20 proverbs in two sets of ten proverbs, each set followed by twelve statements, one of which "explained" each of the ten proverbs, there being two extra statements in each set not explaining the proverbs. The pupils were required to place in the parenthesis before each proverb the number of the statement which explained it.

*Thanks are due to Mrs. Mary D. Chamberlain for the Proverb Test used in this study. The words of the Spelling Test, as explained later, were taken from Ayres' list. (See Ref. 2.)

†Although the Spelling Test was found, in both the preliminary and in the present investigations, to afford quite as good a measure of intelligence, from one point of view, as the other tests (see intercorrelations), it has since been considered best to drop this test from the scale.

The Disarranged Sentence Test consisted of 26 sentences with the words disarranged, as shown in the sample. The pupils were required to rearrange the words mentally to make sense and indicate whether the sentences so constructed were true or false by underlining the words true or false at the end of the line.

The Relation Test consisted of 24 items, each in the form of a proportion in which one of the four terms was to be supplied, indicating by number from five alternative answers given on the same line.

The Geometric Test consisted of 22 items, using as a basic principle that described by Abelson. (See Ref. 1.) Referring to the figures constructed by overlapping one or more circles, triangles, and rectangles, the pupils were required to place figures 1, 2, etc., in certain designated spaces as suggested in the sample.

The Following Directions Test consisted of 14 problems requiring the pupils to place certain numbers in certain figures on the Woodworth and Wells Cancellation Sheet. This test presumably differed from the preceding one in that its difficulty consisted more in the comprehension of involved language while in the Geometric Test, the difficulty lay chiefly in tracing out the space relations.

The Narrative Completion Test was of the type used by Whipple, Ebbinghaus, Terman, and others. It consisted of a short story of which certain words were omitted leaving blanks in which the pupils were required to write the words which in their judgment best fitted into the story.

III. The Preliminary Investigation

To aid in choosing the tests for the Point Scale and in determining the most suitable forms in which to give them, a preliminary investigation was conducted in which 29 pupils in grades 4 to 8 of a small school were tested. Fifteen tests were used. Eight of these were in the same or nearly the same form as those shown in the Appendix. Two others, the Arithmetic and Spelling Tests, have since been entirely made over. The Synonyms Test was given orally first, then about three weeks later repeated in the form shown. The other five tests were: a test in word meaning recognition, of the type suggested by the writer (See Ref. 8) and called the Reading Test; a test in the reproduction, in writing, of sentences dictated, called the Memory for Sentences Test; the Trabue Completion Test (see Ref. 15); the Kansas Silent Reading Test (see Ref. 4); and the Starch Grammar Test (see Ref. 12). The list is shown in Table I. The tests marked with the asterisk were given in duplicate, the double scores being used in the correlations.

The correlation of each test with the composite of all scores except those of the Oral Synonyms and Grammar Tests are shown in the first column of the table. The correlations with mental age as determined by the Stanford Revision of the Binet Scale are shown in the second column. Considering the small number of individuals as well as the unreliability of scores, the coefficients may be regarded as of suggestive value only.

TABLE I.

Some Results of the Preliminary Investigation

Test	Correlation with Composite	Correlation with Mental Age
Relation	.94	.97
Proverbs*	.94	.94
Following Directions	.86	.95
Geometric*	.89	.92
Trabue Completion Test*	.88	.88
Reading*	.92	.82
Kansas Silent Reading Test	.90	.88
Synonyms (Oral)	.79	.87
Synonyms (Written)	.83	.85
Disarranged Sentences*	.86	.81
Narrative Completion	.86	.80
Arithmetic	.84	.80
Spelling*	.79	.84
Memory for Sentences	.77	.82
Memory for Digits*	.72	.42
Starch Grammar Test	.30	.49

Correlation of Mental Age with Composite Score: .94
Same, "corrected for attenuation" (estimate): .99

IV. Acquisition of the Data

Administration of the Tests.—The tests were given to 121 children of a large grammar school—43 in the fourth grade, 40 in the sixth grade, and 38 in the eighth grade. Each test was taken by all of the pupils of one grade at a time in their regular room, the teacher being present. The writer personally conducted all the tests, giving all the directions and explanations. The tests were given in approximately the order indicated in the foregoing section. In most instances two test series were given to each grade each day, one in the morning and one in the early afternoon. The giving of the first and second tests of the same kind were separated by three or more days in most instances but particularly in the cases of the Arithmetic, Geometric, and Following Directions Tests, in

*Double scores used.

which the second test is in the nature of a recast of the first. Time was allowed for all to finish in nearly all cases, except, of course, the Disarranged Sentence Test, which is a speed test. Occasionally when one or two pupils lagged far behind the others, their papers were taken up before they finished. In such cases it was usually noted that the pupils had permitted themselves to be distracted from their work. In the Disarranged Sentence Test sufficient time was allowed for only one pupil to finish. The order, "Stop," was then given and the time noted. For purposes of comparison, all scores were afterward increased to a five minute basis.

The pupils of each grade were adjured at the beginning of the testing not to give or receive aid during the taking of any tests. A wholesome attitude appeared to be taken by all during the testing. In such instances of apparent collusion as were noted, the pupils were quietly cautioned. These instances were few. On the whole the pupils were orderly and attentive and signified their interest in the testing.

Scoring.—In the case of each test except the Synonyms, Spelling, and Disarranged Sentences, one count was given for each correct answer and no count for incorrect or omitted answers. In the case of Synonyms, however, since there are but two alternative answers, S or O, theoretically, of the answers given concerning the pairs of words not known by any pupil, but guessed at, one half will be right by chance. Therefore, if say 35 of the 50 were known and correctly marked, and 10 of the remaining 15 guessed at, leaving 5 blank; of the 10 guessed at, 5 might be marked rightly by chance. This would make 40 correct, 5 incorrect, and 5 blank. It seemed, therefore, that as many counts should be deducted from the total correctly marked (40) as were incorrect (5) thus giving a score of $40-5=35$, the number assumed to be known. A person guessing at all of them and getting half right by chance would then attain a score of $25-25=0$. This method was adopted in scoring the Synonyms and Disarranged Sentence Tests. The case of the Disarranged Sentences is complicated by the fact that sentences wrongly marked on account of haste are penalized additionally by the loss of time in scanning them. The suitability of the method, therefore, should perhaps be investigated.

N = 35

Since there are three possible answers, F. S, or N, in the Spelling Test, theoretically ⅓ of the number of those guessed at would be marked rightly by chance. This would mean that, to follow the

above method, the score should be obtained by deducting from the number rightly marked, ½ the number of those wrongly marked. However, inasmuch as there were only fourteen individuals who did not attempt all the words, and to avoid possible negative scores, the scores were obtained by giving one count for each right answer, no count for each wrong answer, and ⅓ count for each blank, on the assumption that if guessed at, ⅓ of these words would have been rightly marked, This brought all the scores to the same basis and necessitated counting only right answers in all but 14 cases. Identical rank orders of the individuals are obtained from the scores by the two methods. The scores that would have been obtained by using the first method can easily be derived from those used merely by multiplying by 1½ and subtracting 25.

In order to obtain a suggestion as to the value of the above methods of scoring, the sum of the differences between the first and second scores of the 14 pupils above mentioned was found first when the scores were obtained as above and second when obtained by merely counting the number of correct answers, taking no account, therefore, of the element of chance. The sum of the differences in the first case was 34 and in the second 45, although the scores were less in the second case. This suggests that the method employed was the more reliable.

While there is, of course, a 'one-in-five chance' of an element of the Relation Test being marked rightly by guess, it was not deemed necessary to take account of it. In an auxilliary investigation regarding the scoring of the Digit Test, the papers were scored (1) according to the number of digits in the last number correctly reproduced, (2) according to the number of digits in the next to the last number correctly reproduced, and (3) according to the last group of numbers of the same size of which two or more were correctly reproduced. No one of these three methods appeared to be appreciably superior to the others. The reliability coefficient of scores by method (1) was .53, by the method employed in this study, .74. It was discovered after giving the test that some of the pupils were able to reproduce numbers of nine or more digits within three trials. If such had been included in the test, the reliability coefficient by method (1) would no doubt be higher. It is believed, therefore, that with a sufficiently exhaustive test, the loss in reliability of method (1) would be more than made up by the great saving in time of scoring.

Briefly, the plans of scoring were as shown in Table II.

TABLE II.

Summary of Plans of Scoring

Test	Score
Spelling	1 count for each correct answer and ⅓ count for each blank. (nearest whole number)
Arithmetic	1 count for each correct answer.
Synonyms	1 count for each correct answer and 1 count deducted for each incorrect answer (blanks not counted).
Memory for Digits	1 count for each number entirely correct.
Proverbs	1 count for each correct answer.
Disarranged Sentences	1 count for each correct underlining with 1 count deducted for each incorrect underlining.
Relation	1 count for each correct answer
Geometric	1 count for each figure 1 correctly placed, provided no other figure 1 appeared in the same design, and similarly, 1 count for each figure 2 correctly placed.
Following Directions	1 count for each direction correctly followed.
Narrative Completion	1 count for each blank satisfactorily filled.

The scores for each individual in each test will not be given as obtained by the above plan but instead they will be given in an altered form explained below. The scores are given in Appendix II.

V. The Reliability of the Scale

The reliability of a test may be expressed in two ways, either (1) by giving the probable error of a score in the units of the score, the probable error being the value of that error which is exceeded in amount by half the errors, or (2) in terms of the coefficient of correlation between two tests of the same kind. The probable error of a score as a measure of the reliability of a scale is comparable with other values of the probable error found in connection with the testing of other groups of individuals but it is not comparable with the probable error of the scores of other tests unless the units in the two scales measure the same increments of ability. This would not happen often and only accidentally. The reliability coefficient, as has been explained more fully elsewhere (see Ref. 11), derived from measures of one group is not comparable with a reliability coefficient for the same test derived from measures of another group unless the heterogeneity of the two groups is the same or nearly the same. In many instances, of course, this is not the case. The reliability coefficient of one test, however, is comparable with that

of another test when both are derived from measurements of the *same group*. The reliability coefficients are necessary under these conditions to show the relative reliabilities of two tests. They compensate the measures of reliability for inequalities of scale units.

We have therefore found both the probable errors and the reliability coefficients of each of the ten tests. The probable errors of scores in the several tests were found according to the method described at length in Ref. 11. This method is expressed by the formula,

$$P.\,\epsilon. = \frac{\text{Med. Dif.}}{\sqrt{2}}$$

in which Med. Dif. is the median difference between scores by the same individuals in the two tests of the same kind. (A test of the first scale is called Test I; the corresponding test of the second scale Test II.) Before making the subtractions, however, it is necessary to have the scores of both tests in terms of either one or the other of the two tests, since these are quite often somewhat different; due to slight differences in difficulty, to practice effect, etc. For the purpose of evaluating the scores of one test in terms of the other, plots were made in which the scores in Test I were represented as abscissae (horizontally) and those of Test II as ordinates. The manner in which the scores in the two tests corresponded was then found by drawing in each plot a line of relation. This is such a line that the abscissa and ordinate of any point on it represent corresponding scores in the two tests. By inspection of the plots, it was deemed valid to draw a straight line of relation in all cases except that of the Narrative Completion Test, in which it was apparent that the true line of relation was markedly curved. In that case the curve of relation was drawn by the method we have called the method of correspondence by rank (see Ref. 7). In all cases except that of the Narrative Completion Test, the line of relation was obtained by finding the means, M_x and M_y, of the values of x and y and the average deviation, A. $D._x$ and A. $D._y$., of the distributions of values of x and y, and then drawing a line through the point (M_x, M_y) having a slope such that the tangent of the angle formed with the X axis $= \frac{\text{A. D.}_y}{\text{A. D.}_x}$

To find the score in terms of Test II which corresponds to any given score in terms of Test I, it is necessary merely to find the point on the relation line corresponding to the score in Test I and to note the

score in Test II at the left which corresponds to this point. The differences between the scores, in terms of Test II, are measured by the distances of the points of the plot above or below the line; in terms of Test I, by the distances to the right or left. The values of *Med. Dif.* were obtained by the method. (See Ref. 7):

$$\text{Med. Dif.} = .8453\text{x}\text{Avg. Dif.}$$

That is,

$$\text{P. E.} = \frac{.8453\ (\text{Avg. Dif.})}{1.414}$$

The values of the probable errors of each of the several tests were obtained first in terms of Test II and the corresponding values in terms of Test I were derived by dividing by the tangent of the angle of the line of relation. The values of the probable error in both terms are given in Table III.

TABLE III.

Reliability of the Tests

	Probable Errors		Reliability Coefficients	
	Scale I.	Scale II.	Single Tests	Double Tests
1. Spelling	1.49	1.45	.942	.970
2. Arithmetic	.74	.80	.871	.931
3. Synonyms and Antonyms	1.96	1.86	.753	.85)
4. Memory for Digits	1.04	1.21	.746	.855
5. Proverbs	1.17	1.02	.761	.864
6. Disarranged Sentences	1.28	1.76	.737	.849
7. Relation	1.50	1.86	.729	.843
8. Geometric	1.22	1.15	.805	.892
9. Following Directions	.75	.97	.82)	.901
10. Narrative Completion	5.43	3.80	.840	.913
			Total	8.877

The formula used for finding the reliability coefficients was:

$$r = 1 - \tfrac{1}{2}\left(\frac{\text{A. D.}_{(\text{dif's})}}{\text{A. D.}_{(\text{scores})}}\right)^2$$

which is a variation of the difference formula:

$$r = 1 - \tfrac{1}{2}\left(\frac{\sigma_{(y-x_y)}}{\sigma_y}\right)$$

This latter formula is the equivalent of the Pearson product-moment formula. (See Ref. 11.) In these formulae, A. $\text{D.}_{(\text{dif's})}$ and $\sigma_{(y-x_y)}$ are measures of the variability of the distribution of

differences between the scores of each of the 121 pupils in Test I and Test II, when the scores in Test I are evaluated in terms of Test II; and in which A. D.$_{(scores)}$ and σ_y are respectively corresponding measures of the variability of the distribution of *scores* in Test II.

The reliability coefficients thus found for each test are shown in the third column of Table III.

We were quite surprised to find the Spelling Test to be so much in the lead in this rating. However, the Spelling, Narrative Completion, and Synonym Tests had 50 elements while the other tests had only 25 or less. The Arithmetic, Following Directions, and Geometric Tests no doubt have an advantage over the others in that Test II was only slightly different from Test I.

The aim in duplicating the tests, as has been stated, was to make the second test in each case as nearly like the first as possible without actually copying it. This was done in order that the score in the second test would be as near as possible to a second score in the same test. It is possible that a second score in the same test would have been preferable for finding the reliability if it had been convenient to separate the two givings of the tests by a sufficient interval. Even this, however, would introduce new sources of error. Since the differences in difficulty between the two tests of a kind are not the same for all the pupils, the differences between the scores in the two tests tend to be greater than would be the case it the same test could be given twice, even without memory of the first testing, in which case the difference in the scores would be due merely to differences in disposition at the times of taking the first and second tests. For this reason, the values of the probable errors and reliability coefficients, considering only errors due to varying disposition, are really less than those given here.

Further consideration regarding reliability will be given later. These depend upon the values of inter-test correlations.

VI. Graduation of the Scale

Theoretical Considerations.—There are two aspects to the graduation of the scale. One deals with the proper combining of the scores of the several tests and the other with the finding of age norms, percentage norms, etc. The scores of each individual in the ten tests must first be combined into a single score, say a "point-score," and then those point-scores may be determined which are

normal for each of the given ages of childhood, or those which given percentages of adults may be expected to attain, etc.

In order that the scores of an individual in the several tests may be properly averaged, it is necessary to take account of the differences in value of the units of the scales of the several tests. If an increment of one problem in an Arithmetic score is in reality equal to an increment of four words in the score of the Synonym Test, to average the scores in the two tests just as they stand would be to give the Synonym Test four times as much weight as the Arithmetic Test. If, therefore, it is desired to give equal weight to each test, the score of an individual in each test must be transmuted into other terms, say "points," such that *equal increments of ability* in each test receive *equal increments of points.* It is convenient, also, while assigning point values to the scores in the several tests, to arrange that correspondihg *amounts* of ability in the several tests shall receive corresponding *numbers* of points. The first of these conditions is essential and the second convenient for the purpose of averaging scores properly; both conditions are essential for the purpose of comparing scores in the several tests with one another. If it seemed reasonable to assume that for the individuals of a given group, the ability possessed by the upper 25% in any test was as much above that possessed by the upper 50% as that ability was above the ability possessed by the upper 75%, and if the ability in any one test was considered equal to that in any other test which was possessed by the same percentage of individuals, then the first of the above mentioned conditions would be complied with by representing the difference between upper 25% and upper 50% ability in each of the several tests by some number of points (say 10), and the difference between upper 50% and upper 75% ability in each of the several tests by that *same* number of points (10). And the second condition would be complied with by representing 50% ability in all of the ten tests by the same number of points (say 50) in which case, of course, 25% ability would be represented in each case by 60 points and 75% ability by 40 points.

We have been speaking of the equality of increments of ability, but such equality is a very indefinite thing. Equal increments of ability must be such as are measured by the same number of units of some kind. We have not been willing to grant that the steps of any test scale necessarily measured equal increments of ability.

Nor would we admit that any year's growth in ability is equal to every other year's growth. The growth of ability is supposed to retard eventually with age. In what units then will we say ability may be measured so that equal numbers of units measure equal increments of ability? In a previous article (Ref. 10) we have suggested that absolute units of ability be so defined that the distribution of abilities of all adults *will be normal* (in the technical sense). This would mean that those percentages of adults which were considered as possessing abilities which marked successive steps on an absolute scale of ability were the same percentages as those of the normal probability surface which corresponded to successive units of the base. Until such time, however, as a very large number of unselected adults have been tested, such a criterion of equality of units of ability will be unavailable. In lieu of such a criterion, an alternative method was used.

The Procedure Used for Determining Equality of Increments of Ability.—Although we have felt that the units in one part of a single test scale were very apt to be of greater value than those in some other part, it is quite probable that if the upper units of some test scale must be considered as measuring greater increments of ability than the lower units, the opposite probably might be considered true of some other test scale, so that taking the test scales all together, the median value of the units in one part may be considered as equal to the median value of the units in any other part. Proceeding upon that hypothesis, the most probable true form of the distribution of abilities of the 121 pupils was determined by obtaining a composite of the separate distributions for the ten tests as follows:

1. The score attained in each test by the 30th individual in rank (beginning with the lowest) was assigned a preliminary point-value of 40 points and the score attained by the 90th individual in rank was assigned a preliminary point-value of 60 points.*

2. Tentative point values corresponding to all the other scores were then determined in such a manner that the units in all parts of the test scale were represented by equal increments of points. This was accomplished graphically in each case by drawing a straight line.

3. From the smooth curves of distribution of test scores were

*These scores were not the *actual* scores of those individuals but the scores corresponding to them on smooth curves through the distributions of consecutive scores.

then determined the scores attained by the 3rd, 9th, 15th, 60th, 105, 111th, and 117th individuals in rank order. These points in the distribution curves were believed to best reveal any skewness of the distribution.

4. The preliminary point values corresponding to the scores attained by the 3rd individual in each test distribution were then ascertained. These were then plotted in order of magnitude and a median value determined by means of a smooth curve through the plotted points. This median point value was 24.4. The other median point values were as follows:

Individual in order:	3,	9,	15,	(30)	60,	(90)	105,	111,	117
Point value	24.4	29.7	33.3	(40)	50.1	(60)	66.7	70.1	75

It should be stated that these values indicate that the distribution of abilities of the 121 pupils approximately normal.

5. Since the median of the preliminary point values obtained by the 3rd individual in rank in the several test distribtuions was 24.4, this value may be assumed to be the most probable true value, in terms of our established absolute units, of the ability in any test which the 3rd individual in rank order attained. The score in each test attained by the 3rd individual in rank order (by the curve) was then given, therefore, the corrected point value 24.4. Similarly the score in each test attained by the 9th individual was then given the corrected point-value 29.7, etc.

6. In order to determine the corrected point value to be similarly assigned to all the other scores in each test, a graph was made for each test in which the preliminary point values corresponding to the scores attained by the 3rd, 9th, etc., individuals were plotted as ordinates and the new point values, 24.4, 29.7, etc., plotted as abscissae. A smooth curve was then drawn through the series of plotted points. This curve was then taken as showing the relation between the preliminary and corrected point values corresponding to each score in the test From this curve for each test were taken the corrected point values corresponding to each score. These are shown in Table IV. They no doubt represent the nearest approach that can be made to a true absolute point scale.

Considerations with Regard to Weighting and Combining the Scores. —After finding the corrected point values corresponding to each test score, the scores of each pupil in each test were transmuted into terms of points and the total score found for each. These are given in Appendix II.

This method of combining the scores resulted in equal weight being given to each test. No doubt some of the tests are more significant than others in the measurement of general ability, however we conceive it. Unreliability of a test, of course, lowers its significance. Other aspects of significance depend upon the conception of general ability. If a test is considered as measuring general ability only to the extent to which the factors entering into the ability tested are common to other abilities, both as to number of factors and as to number of abilities to which they are common, then the degree to which a test may be considered as measuring

TABLE IV

Showing the Number of Points Corresponding to Each Score in Each Test

Score	Spelling Points	Arith. metic Points	Synon- yms Points	Digits Points	Proverbs Points	Disar. Sentences Points	Relation Points	Geomet. Points	Fol. Direc. Points	Narra. Comple- tion Points
0		21		20	32	23	25	21	23	26
1		24		20	34	28	26	22	27	26
2		27		21	35	33	27	24	31	27
3		30		22	36	38	28	25	35	27
4		33		23	37	43	29	27	39	28
5		36		24	39	47	31	28	43	28
6		39		26	40	50	32	29	48	29
7		42		28	41	53	33	31	52	29
8		45		31	43	56	35	32	56	30
9	20	48		35	44	58	37	34	60	30
10	21	51		38	45	61	40	36	64	31
11	22	54		42	47	63	42	38	68	31
12	23	56		45	48	66	45	41	72	32
13	24	59		48	49	68	47	43	76	32
14	25	62		51	50	71	50	46	80	33
15	26	65		55	52	73	52	49		33
16	26	68		58	53		55	52		34
17	27			61	55		57	56		34
18	28		20	64	56		60	59		35
19	29		22	68	57		62	63		35
20	30		23	71	59		65	66		36
21	31		25				67	70		36
22	32		26				70	74		37
23	32		28				72			38
24	33		29				75			38
25	34		31				77			39
26	35		32							39
27	36		34							40
28	37		35							41
29	38		37							41
30	38		38							42
31	39		40							43
32	40		41							44
33	41		43							45
34	41		44							46
35	42		46							47
36	43		47							48
37	44		49							49
38	45		50							50
39	46		51							51
40	47		53							52
41	48		54							53
42	49		56							55
43	50		57							56
44	51		59							57
45	52		60							58
46	53		62							59
47	55		63							61
48	56		65							62
49	58		66							63
50	60		68							64

general ability is expressed by the amount of "correlational spread" of the test, to use McCall's expression, by which is meant the sum of the intercorrelations of the test with other tests comprising a fairly representative collection, each presumed to involve factors common to the others. The last qualification is necessary since, if the group of tests is too restricted in kind, certain 'specific' abilities may be common to too large a proportion of the tests and thus vitiate the criterion of general ability.*

On the other hand if a test is considered as contributing to the measure of general ability if it measures an ability that may be considered valuable in aiding the individual to adjust himself to the new problems and conditions of life, whether such ability has few or many factors in common with others; then it is not proper to use only the criterion of correlational spread. Two possible alternatives suggest themselves. If there were available for the individuals tested a satisfactory criterion of their powers of adaptation to the new conditions and problems of life, in the nature of a measure of economic or scholastic success, then it would be necessary merely to weight the tests according to the regression equation method, so as to obtain the best correlation of the composite score with the criterion. In lieu of such a criterion, the tests might be weighted according to a combination of the weights assigned by a number of judges. In this study, for instance, the results of all the tests except that of Memory for Digits correlated uniformly highly with each other. The Digit Test, which showed a reliability not the least among the ten tests, stood quite apart from the other tests in showing low correlations with all of them. According to the criterion of correlational spread, this test would be weighted very much lower than any of the others. According to either of the criteria pertaining to the second conception of general ability, however, the Digit Test might perhaps deserve a weight more nearly the amount of the others.

---

*Some mathematical reasoning bearing on this point is given in Appendix III, 1 and 2.

McCall used this criterion in his study (Ref. 5). Another criterion which he also used was the correlation of each test with "Composite," a measure obtained by combining the scores of all the tests (with some exceptions) after weighting each according to *a priori* considerations as to the value of the tests. Although the correlations of the several tests with Composite appear to have been determined by McCall by separate calculations, it would have been possible to obtain the values of these correlations with Composite more simply from the values of the inter-test correlations. The necessary procedure is given in Appendix III, 3.

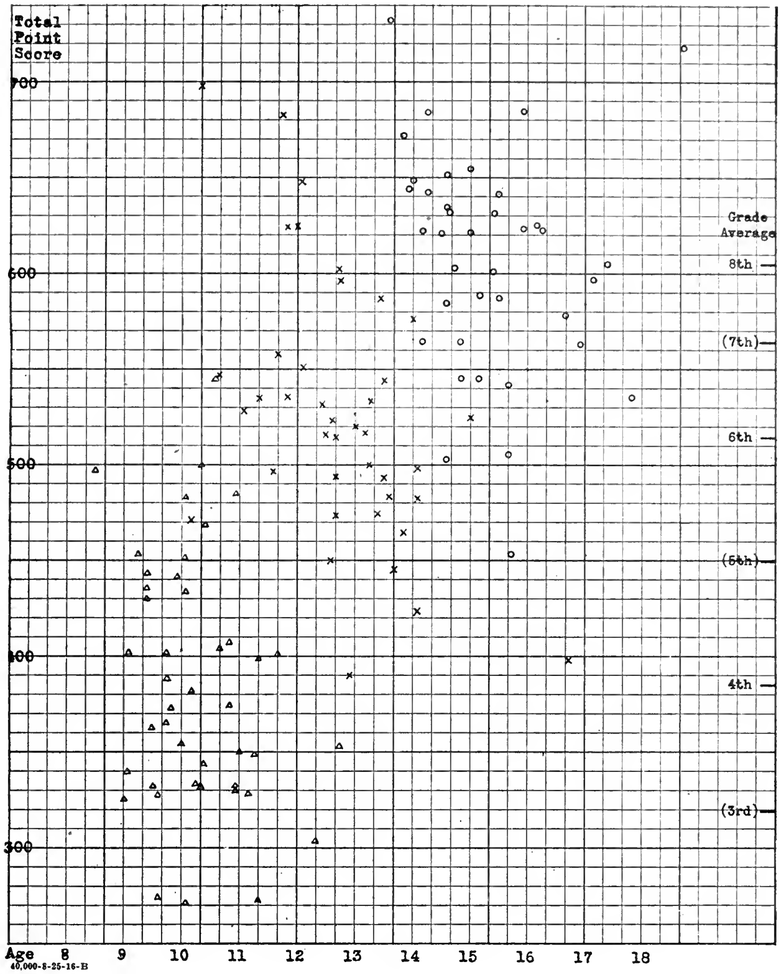


FIG. 1

Showing the Relation between Total Point Score and Age

In this study we are inclined more to the second conception of general ability mentioned. It was not feasible in this study, however, to use either of the criteria appropriate to this conception. To weight the tests according to reliability alone, it would be necessary to weight each inversely in proportion to the square of the probable error (the probable errors being in comparable terms). (See Merriman, Ref. 6, p. 95.) Such procedure, however, prac-

tically implies that all the tests aim to measure the same thing. But since they do not, any weighting given to compensate for different degrees of reliability, necessarily also emphasizes the effect of certain particular abilities and is to that extent undesirable.

For these reasons we have combined the test scores without weighting them.

Finding Age Norms in Terms of Point Scores.—For finding age norms, a plot was made. (See Fig. 1.) One point pertains to each pupil. The abscissa of each point represents the pupil's age and the ordinate his total point score. In order to find the score which would be considered normal for 10-year-olds, the average score was found of all pupils of ages from 9 years, no months, to and including 11 years, no months; for 11-year-olds, the average score was found of all pupils of ages 10 years to and including 12 years, etc. The norms thus found were as shown in Table V. These values were then plotted. (See Fig. 2.) To our surprise, the points representing the norms for ages 10 to 14 lay in almost a perfectly straight line, which suggests that they are fairly reliable, at least, for the school population tested. This was not expected considering the gaps left by omitting the fifth and seventh grades from the group tested. The norms for years 15, 16, 17 may be seen to fall below the line, the latter two quite markedly. This was to be expected, of course, since the pupils of these ages were selected, being retarded in their schooling. While the true norms for these ages are doubtless above the average values obtained, it was not deemed proper to continue the straight line. The line was therefore curved off to the right as shown. We must regard the norms for the ages above 15, as being only roughly approximate.

TABLE V.

Showing Age Norms in Point Scores

Age:		8	9	10	11	12	13	14	15	16	17	18	19
Point Score Norms:	Observed:			404	446	487	527	566	583	550	584		
	Smoothed:	324	364	405	445	486	526	566	600	624	638	647	650

Completing the Absolute Point Scale.—We have previously (see Ref. 10) given the name, Coefficient of Brightness, to the quotient that would be obtained by dividing the measure of the absolute amount of mental ability of any individual by the measure of the absolute amount of mental ability which was normal for the age of that individual. This means, of course, that the measures of

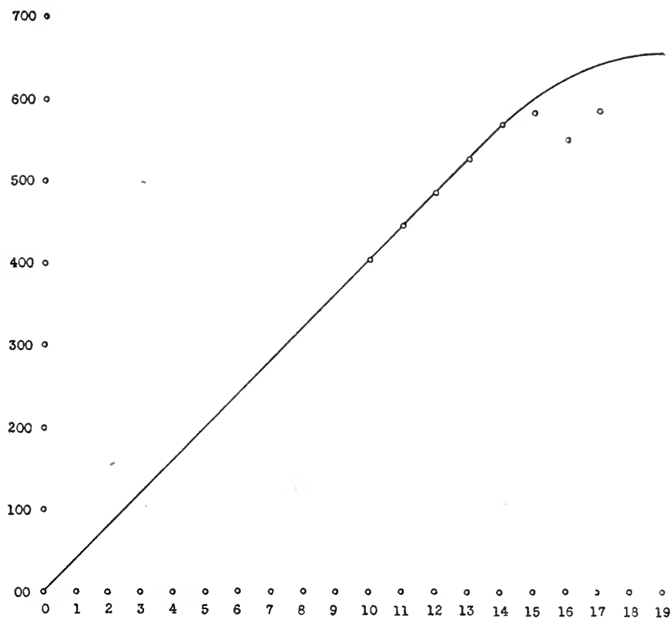


FIG. 2
Showing a Smooth Curve through the Age Norms of Total Point Scores

mental ability must be in such terms that not only will equal increments of ability be measured by equal increments of the scale, but twice as many units on the scale will represent twice as much ability, etc. In other words, zero of the scale must represent just absence of ability. Before it was possible for us to find the coefficients of brightness of the pupils tested in this case, therefore, it was required to note what correction was necessary in the scale of points in order that the number of points representing the ability of age 0 would be 0. The ages for which we may presume to have obtained fairly reliable norms are only those from 10 to 14. Inasmuch, however as the increments of points between the norms for these ages are almost exactly the same, it was regarded as proper to assume for present purposes that if continued, the line through the norms would be straight the rest of the way to age zero. It was then necessary to note what number of points thus corresponded to age zero, this number to be considered the true absolute zero of the final point scale. To our further surprise, it was discovered that by calling the yearly increment of points (below 14) approx-

imately 40.5 a line, which would pass as nearly as any other through the five norms, actually reached zero age at zero of the point scale. This, of course, was an entirely accidental coincidence and not at all necessary. It merely saved us the obligation of subtracting or adding a constant to each of the corrected point values assigned to the several test scores in order to obtain the final point values constituting the completed Absolute Point Scale.

The Determination of the Coefficients of Brightness.—Since the point values in which the scores of the pupils were expressed proved to be those of the Absolute Point Scale, in order to find the coefficients of brightness of each pupil, it was necessary merely to divide the total point score of each by the score which was normal for his age. The norms for the fractional ages were taken from the curve in Fig. 2. The coefficients of brightness thus found are given in Appendix II.

Appendix I.

Sample Extracts of Tests:

Test 1: Spelling

1. forenoon	fournoon	(F)
2. intrest	interest	(S)
3. neighber	neighbor	( )
4. concider	consider	( )
5. entertain	entertane	( )
etc.	etc.	etc.

Test 2: Arithmetic

1. If a boy has 10 cents and then earned 5 cents, how much did he have then? . . . . ( ) cents

7. How many years will it take a glacier to move 1000 feet at the rate of 100 feet a year? . . . . ( ) years

15. A ship has provision to last her crew of 50 men 6 months. How ong would it last 30 men? . . . . ( ) months

Test 3: Synonyms and Antonyms

1. large	big	(S)
2. decrease	increase	(O)
3. empty	vacant	( )
4. knowledge	ignorance	( )
50. conservative	radical	( )

Test 4: Memory for Digits

1. 4739	( ) ( ) ( ) ( )
2. 2854	( ) ( ) ( ) ( )
3. 7261	( ) ( ) ( ) ( )
4. 31759	( ) ( ) ( ) ( ) ( )
5. 42385	( ) ( ) ( ) ( ) ( )
6. 98157	( ) ( ) ( ) ( ) ( )

Test 5: Proverbs

Pr verbs

(3) Make hay while the sun shines.
(   ) In a calm sea every man is a pilot.

Statements to explain the proverbs.

1. Deeds show the man.
2. Leadership is easy when all goes well.
3. Make the best of your opportunities.

Test 6: Disarranged Sentences

1. name a John is boy's..(*true* false)
2. sun morning the the in sets.................................(true *false*)
3. trees birds nests the in build...............................(true false)

Test 7: Relations

hand : arm : : foot : (   )
2. hat : head : : thimble : (   )
23. education : ignorance : : (   ) : poverty

1. 1 leg, 2 toe, 3 finger, 4 wrist, 5 elbow.
2. 1 finger, 2 needle, 3 thread, 4 hand, 5 sewing.
23. 1 laziness, 2 school, 3 wealth, 4 charity, 5 teacher.

Test 8: Geometric Test

(*Designs were presented composed of two or more geometrical figures—circles, rectangles, and triangles—overlapped.*)

1. Place a figure 1 so that it will be both in the rectangle and in the circle.
7. Place a figure 1 so that it will be in both circles, in the triangle, and in only one rectangle.

Test 9: Following Directions

(*A page of Woodworth and Well's Cancellation Test was supplied each pupil.*)

2. In line 1 [of the forms] place a figure 1 in the first star and a figure 2 in the second circle.
7. In line 5, place a figure 7 in the form which follows the same kind of form as that which follows it.

Test 10: Narrative Completion

Once upon a .... there was a y................ .... who was very p........ . He went from place to .... trying to find ....

Appendix II.

Showing the Point-Scores of each individual in each test.

Eighth Grade

Pupil	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16
Spell.	77	55	68	56	63	64	59	62	65	57	54	70	72	53	63
Arith.	76	54	67	49	51	76	56	58	73	56	72	60	64	44	62
Synon.	76	58	64	50	57	67	72	59	57	60	56	75	69	47	64
Digit.	80	76	41	45	43	62	53	50	41	50	62	32	45	30	36
Proverb.	81	54	65	35	46	62	65	54	68	60	52	68	71	35	62
D. Sen.	68	72	63	63	68	75	63	72	60	59	71	53	72	59	63
Rel'n.	61	53	71	44	53	49	69	69	67	69	67	75	73	41	63
Geom.	52	55	58	55	58	58	66	69	74	66	64	58	61	45	61
Fol. D.	67	58	65	48	56	51	67	61	58	72	58	75	75	43	56
Compl.	80	62	64	58	68	58	65	68	68	72	48	68	83	56	58
Sums	718	597	626	503	563	622	635	622	631	621	604	634	685	453	588
C. B.	109	93	100	86	89	108	108	99	103	103	102	108	110	74	96

Pupil	17	18	19	20	22	23	24	25	26	27	28	29	30	32	33
Spell.	64	70	67	65	51	64	56	64	54	72	65	72	61	54	56
Arith.	66	58	49	64	58	64	70	56	66	64	70	56	64	73	51
Synon.	58	67	65	64	56	61	55	54	65	68	68	59	54	66	65
Digit.	48	62	50	60	45	43	30	57	76	57	41	66	30	64	57
Provb.	52	71	60	65	52	52	52	54	60	81	65	58	56	68	50
D. Sen.	67	63	60	70	68	64	53	73	65	63	68	66	71	75	68
Rel'n.	67	61	55	75	46	61	55	53	46	61	71	42	60	69	65
Geom.	55	69	52	64	64	40	49	66	66	61	64	61	64	55	72
Fol. D.	58	67	38	58	53	53	58	53	63	73	61	61	48	65	73
Compl.	54	63	50	56	43	63	64	55	44	72	69	60	57	66	66
Sums	589	651	546	641	536	565	542	585	605	672	642	601	565	655	623
C. B.	98	111	91	104	84	98	88	99	95	120	111	99	95	109	100

Fourth Grade

Pupil	34	35	36	37	38	39	40	41	51	52	53	55	56	58	59
Spell.	64	64	64	54	68	51	70	64	27	32	39	31	53	46	29
Arith.	58	60	67	77	70	67	70	79	44	27	39	34	19	46	39
Synon.	50	53	74	48	72	63	46	73	19	35	32	32	25	43	29
Digit.	55	45	75	75	53	36	64	78	41	36	68	50	45	50	36
Provb.	46	50	68	56	56	52	71	75	40	31	35	24	31	31	31
D. Sen.	49	59	68	70	68	65	65	63	48	38	40	25	48	49	42
Rel'n.	58	46	61	56	60	75	75	82	26	32	46	29	30	38	38
Geom.	29	52	72	66	61	58	66	72	29	36	41	43	25	45	36
Fol. D.	43	61	72	53	70	51	58	75	29	29	46	41	29	41	32
Compl.	54	56	62	66	66	60	64	71	27	36	44	35	35	52	38
Sums	506	546	683	621	644	578	649	732	330	332	430	344	340	441	350
C. B.	83	92	119	107	115	92	115	133	75	87	113	82	93	110	79

Pupil	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74
Spell.	36	40	40	43	30	41	46	53	37	29	31	35	39	38	37
Arith.	27	23	39	32	36	32	39	49	36	23	44	36	44	46	32
Synon.	43	30	29	37	32	47	62	48	50	31	29	27	40	54	40
Digit.	57	27	34	41	55	57	48	45	57	50	50	53	43	69	30
Provb.	35	31	31	24	35	48	68	56	40	31	31	24	45	50	40
D. Sen.	38	27	32	37	41	35	46	41	45	27	40	40	40	34	34
Rel'n.	44	28	25	31	12	55	44	44	33	25	49	38	56	51	34
Geom.	45	22	38	22	28	49	55	55	23	30	45	41	58	45	38
Fol. D.	32	21	38	29	35	46	48	48	38	35	41	29	43	51	43
Compl.	31	22	23	30	24	59	41	61	45	22	39	26	43	47	38
Sums	388	271	329	326	328	469	497	500	404	303	399	349	451	485	366
C. B.	99	67	73	90	85	112	145	120	91	61	87	77	111	110	93

Pupil	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
Spell.	45	32	42	40	36	50	33	46	35	34	44	29	44	51	15
Arith.	41	27	27	32	36	46	36	32	44	39	34	36	49	49	34
Synon.	50	34	27	28	43	51	28	30	49	33	28	39	36	54	46
Digit.	34	24	64	64	45	57	39	39	64	55	22	22	43	45	50
Provb.	48	38	24	35	31	43	40	43	52	31	38	40	38	35	48
D. Sen.	45	25	28	30	43	54	32	45	45	38	23	32	40	45	44
Rel'n.	42	36	32	34	41	53	31	46	55	39	32	33	38	42	29
Geom.	61	52	25	38	33	64	38	41	45	26	43	45	52	41	37
Fol. D.	43	29	29	43	46	61	41	41	51	35	29	46	51	32	35
Compl.	34	37	33	37	48	66	35	39	42	33	39	33	43	42	36
Sums	443	334	331	381	402	545	353	402	482	363	332	355	434	436	374
C. B.	117	80	75	93	102	127	68	110	119	95	80	88	107	115	94

Sixth Grade

Pupil	90	91	92	94	95	96	101	102	103	104	105	107	108	109	110
Spell.	42	8	34	28	33	31	35	50	52	54	52	49	58	40	64
Arith.	41	36	34	32	32	51	62	62	54	44	49	44	49	49	51
Synon.	48	22	43	22	37	34	74	56	45	45	54	50	59	46	59
Digit.	60	30	50	29	39	73	50	60	55	24	69	29	50	62	48
Provb.	40	24	35	35	35	43	65	62	43	40	62	38	45	43	52
D. Sen.	45	41	51	19	45	32	47	40	48	40	41	56	53	45	47
Rel'n.	44	29	41	31	35	38	55	60	58	42	60	53	46	38	58
Geom.	43	28	32	24	49	41	45	66	55	31	45	45	52	38	49
Fol. D.	41	32	43	29	38	32	58	75	56	25	56	43	61	46	51
Compl.	49	23	38	26	32	33	44	71	59	45	45	39	72	58	50
Sums	453	273	401	275	375	408	535	602	525	390	533	446	545	465	529
C. B.	122	59	85	71	85	93	117	117	87	75	99	80	99	83	118

Pupil	111	112	114	115	118	119	120	121	122	123	124	125	127	128	129
Spell.	55	44	61	47	51	70	36	41	56	47	51	59	56	68	56
Arith.	51	44	54	39	51	56	56	46	41	44	51	58	79	49	49
Synon.	49	40	62	51	41	75	43	50	55	54	53	45	72	51	52
Digit.	62	45	69	50	27	75	57	55	41	32	53	45	57	68	60
Provb.	58	56	75	56	38	71	56	56	50	45	48	58	65	46	52
D. Sen.	40	45	49	56	51	55	33	41	61	45	48	26	56	47	48
Rel'n.	44	53	58	48	23	69	48	55	46	56	60	71	73	58	55
Geom.	45	38	81	49	41	77	55	43	47	52	47	69	81	61	49
Fol. D.	48	48	70	38	32	78	56	61	46	51	58	53	72	51	65
Compl.	45	37	69	39	44	72	54	52	56	45	48	48	61	52	58
Sums	497	450	648	473	399	698	494	500	499	471	517	532	672	551	544
C. B.	106	88	132	92	64	168	96	93	87	115	102	106	144	112	127

Pupil	130	131	132	133	134	135	136	138	139	140	141	142	143	144	145	146
Spell.	64	55	53	45	64	46	52	47	44	47	51	49	40	52	61	59
Arith.	49	64	69	60	41	39	51	54	58	46	39	51	62	67	51	56
Synon.	73	62	62	55	48	43	48	43	53	55	56	55	52	48	63	64
Digit.	43	50	34	64	60	53	69	34	75	62	48	53	30	53	48	62
Provb.	81	65	62	58	54	40	40	52	60	65	60	58	46	45	71	71
D. Sen.	54	58	52	46	49	38	49	45	41	65	44	43	46	49	52	57
Rel'n.	61	60	55	58	48	39	48	56	60	55	49	36	58	58	49	67
Geom.	72	69	52	52	49	38	58	66	69	35	40	41	49	66	52	49
Fol. D.	65	51	56	43	58	43	61	41	58	43	53	38	51	38	70	80
Compl.	63	63	63	55	53	44	44	46	60	42	53	51	48	41	70	60
Sums	625	597	558	536	524	423	520	484	578	515	493	475	482	517	587	625
C. B.	131	115	118	112	102	74	99	88	101	100	90	87	85	97	108	129

Concluded in the June Number

LOGICAL MEMORY AND SCHOOL GRADES

IRVING KING AND T. B. HOMAN

University of Iowa

It was the purpose of this study to determine what sorts of tests of immediate memory for sense material show the best if any correlations with scholastic marks. But few investigations of this problem have been reported and the results are fragmentary and conflicting. The most extended report is by A. Travis, in *The Psychological Clinic* of December, 1915, "Reproduction of Two Short Prose Passages: A Study of Two Binet Tests." Travis found that these two short prose passages from the Binet tests, when given to groups of college students, and immediate reproduction is called for, not only do not correlate in the results with one another but do not show any correlation with class marks. On the basis of these results Travis raises the question as to whether other memory tests of a similar character but of different material might be devised which *would* correlate with other abilities. The writers of this article undertook a further investigation of this question. They sought to determine whether the *length* of the prose passage used as a test is of material importance for the test and whether the *kind* of material used is an important element. The different types of material selected may be roughly described as *historical, descriptive,* and *narrative.* Each type of material was represented by tests of three different lengths, a short passage which required from one to one and one-half minutes for the experimenter to read, a medium length passage, requiring from two to three minutes for the reading and a slightly longer passage requiring from four to six minutes for the reading. Special effort was made to secure passages in which the thought was clear-cut. Copies of the passages used will be furnished any one desiring them. For the sake of brevity only the results are reported here.

The subjects available for these tests were 64 pupils in the 5th to the 8th grades of the University of Iowa Elementary School, 72 pupils in the University High School and 110 Junior and Senior students in the College of Liberal Arts.

A different set of paragraphs was used with the High School and University students from what was used with the pupils of the elementary school but all groups had the *short, medium* and *long* tests of the three types, except that the College group took only the narrative paragraphs. The elementary and high school groups therefore each took nine *immediate* memory tests and the college group three. Each group took also the same number of *deferred* memory tests, being called upon to give a second reproduction of each passage after a lapse of from 24 to 48 hours. The results of the deferred reproductions were also compared with scholastic standings.

These tests were all given by the regular teachers to the various class groups which, in most cases, consisted of about 35 pupils or students each, Definite and uniform instructions for conducting the tests were followed by all the teachers. After the first day, before a new paragraph was presented, the subjects wrote what they could recall of the preceding test paragraph.

In addition to the scholastic marks which were to be available for these groups, a series of "mental tests" marks were secured for the high school groups. These pupils, early in the first semester of the year 1916–17 were given a part of the tests used with the University Freshmen, viz: Courtis Arithmetic, Series B, three paragraphs of Simpson's Completion test; a set of twenty harder mixed relations or analogies; the three inch cube test; an opposites test (40 of the easier words in Simpson's hard lists; see the standardizations reported by Gold and King, Jour. Ed. Psych., October, 1916); and logical memory, (The Dutch Homestead). We therefore had two separate measures of ability for the high school group.

All the memory tests were graded on the basis of thought units of the original passages reproduced, full credit being given for all reproduced thought units whether in the same or in different words.

All correlations were computed on the basis of ranks in scholarship and in tests, the Pearson formula adapted to rank differences being used.

Results with Elementary Subjects

The following table gives the correlations computed between the various immediate and deferred memory tests and the scholastic ranks of the elementary school group.

TABLE I.

CORRELATION TABLE — ELEMENTARY GROUP

Showing Coefficients of Correlation Between the Several Memory Tests and the Scholastic Rank

	"r's"
Scholastic ranks and His. Short Im.	.35
" " " " " Def.	.37
" " " " Med. Im.	.36
" " " " " Def.	.55
" " " " Long Im.	.39
" " " " " Def.	.44
Average	.41
Average of immediate memory test correlations	.37
Average of deferred memory test correlations	.45
Scholastic ranks and Desc. Short Im.	.52
" " " " " Def.	.48
" " " " Med. Im.	.34
" " " " " Def.	.32
" " " " Long Im.	.29
" " " " " Def.	.38
Average	.42
Average of immediate memory correlations	.40
Average of deferred memory correlations	.43
Scholastic ranks and Nar. Short Im.	.37
" " " " " Def.	.36
" " " " Med. Im.	.12
" " " " " Def.	.47
" " " " Long Im.	.38
" " " " " Def.	.35
Average	.24
Average of immediate memory correlations	.29
Average of deferred memory correlations	.39

Average of all Short	.42
" " " Medium	.37
" " " Long	.37

It will be noted from this table that the average correlations of the second reproductions in each of the three groups of passages are slightly higher than the correlations of the immediate reproductions. The tests with the descriptive material yield slightly better correlations than do the other groups of material and the short passages yield a higher average correlation than do the medium and long.

Results with High School Subjects

The ranks of the high school group based on the combined mental tests were correlated with scholastic ranks, yielding an r = .54.

Table II gives the correlations between "mental test" ranks, scholastic ranks and memory ranks.

TABLE II.

High School Group — Correlation Table

Showing Coefficients of Correlation Between the Logical Memory Tests and the Combined Intelligence Rank

				"r's"
History	Short	Im.	Mental test rank	.42
"	"	Def.	" " "	.41
"	Med.	Im.	" " "	.29
"	"	Def.	" " "	.09
"	Long	Im.	" " "	.32
"	"	Def.	" " "	.31
			Average	.30
Average of immediate memory correlations				.34
Average of deferred memory correlations				.27
Descriptive	Short	Im.	Mental test rank	.30
"	Def.		" " "	.25
"	Med.	Im.	" " "	.5
"	"	Def.	" " "	.54
"	Long	Im.	" " "	.34
"	"	Def.	" " "	.43
			Average	.40
Average of immediate memory correlations				.39
Average of deferred memory correlations				.40
Narrative	Short	Im.	Mental test rank	.23
"	"	Def.	" " "	.33
"	Med.	Im.	" " "	.30
"	"	Def.	" " "	.36
"	Long	Im.	" " "	.41
"	"	Def.	" " "	.36
			Average	.33
Average of immediate memory correlations				.31
Average of deferred memory correlations				.35

Average for all Short	.32
" " " Medium	.35
" " " Long	.36

Here as in the case of the Elementary School group the results from the descriptive passages yield the highest correlations, and in two of the sets the second reproduction gives better results than the first.

Table III gives the correlations of the various tests with average scholastic ranks of the High School group.

TABLE III.

HIGH SCHOOL GROUP CORRELATION TABLE

Showing Coefficients of Correlation Between the Logical Memory Tests And the General Scholastic Rank

				"r's."
History	Short	Im.	Scholastic ranks	.27
"	"	Def.	" "	.20
"	Med.	Im.	" "	.27
"	"	Def.	" "	.21
"	Long	Im.	" "	.19
"	"	Def.	" "	.29
			Average	.23
Average of immediate memory correlations				.29
Average of deferred memory correlations				.32
Descriptive	Short	Im.	Scholastic ranks	.12
"	"	Def.	" "	.09
"	Med.	Im.	" "	.55
"	"	Def.	" "	.44
"	Long	Im.	" "	.20
"	"	Def.	" "	.45
			Average	.31
Average of immediate memory correlations				.29
Average of deferred memory correlations				.32
Narrative	Short	Im.	Scholastic ranks	.15
"	"	Def.	" "	Neg.
"	Med.	Im.	" "	.22
"	"	Def.	" "	.25
"	Long	Im.	" "	.34
"	"	Def.	" "	.29
			Average	.19
Average of immediate memory correlations				.22
Average of deferred memory correlations				.16
Average for all Short				.13
" " " Medium				.32
" " " Long				.29

Table IV gives the correlations of the various tests taken by the college group with their average scholastic ranks.

TABLE IV.

CORRELATION OF LOGICAL MEMORY WITH SCHOLASTIC RANK. CORRELATION TABLE

Juniors and Seniors, College of Liberal Arts

					"r's"
Narrative	Short	Im.	Scholastic	rank	.06
"	"	Def.	"	"	.08
"	Med.	Im.	"	"	.47
"	"	Def.	"	"	.43
"	Long	Im.	"	"	.43
"	"	Def.	"	"	.45
				Average	.32

Average of immediate memory correlations........ .32
Average of deferred memory correlations........... .32

From these tables it seems that the longer passages are better tests than the shorter ones, the coefficients of correlation between the grades in these tests and scholastic standings being fairly high. Other comparisons of the tests with the different groups of subjects may be easily made by inspecting the tables. The most significant fact is perhaps that nearly all the correlations are fairly high and that only one is negative. On the whole the medium length passages yield the highest average correlations for all groups taken together, namely, .38. The average correlation of the short passages was .29 and of the long was .35.

The average correlation for the historical passages was .32; for for the descriptive .38 and for the narrative, .34.

In Table V is given the median percentage of ideas remembered by the different scholastic groups.

TABLE V.

Median Percentages of Ideas Reproduced (immediate recall) by the Different Groups

GROUP	MEDIAN PERCENTAGE OF IDEAS REMEMBERED
Elementary	25.
High School	32.
College, L. A	39.

Table VI gives the percentages of subjects in the different groups who had higher average standings in the deferred recall than in the immediate.

TABLE VI.

Percentages of Subjects in the Different Groups Who Had Higher Average Results in the Deferred Recall than in the Immediate

Group	Percentage of Subjects Who Gained in the Deferred Recall
Elementary	19.
High School	40.
College, L. A.	45.

General Conclusions

1st. As the pupils advance in scholastic grade their power to remember and immediately reproduce prose passages of the type here used increases.

2nd. The pupils' ability to reproduce such prose passages after an interval of 24 to 48 hours increases more rapdily with scholastic grade than does their ability to reproduce them immediately.

3rd. All coefficients of correlation obtained in these various processes were positive with the exception of the one between logical memory, narrative type, short selection, deferred recall and the general ability of the high school group.

4th. The elementary group showed the highest coefficient of correlation between the general scholastic ranks and the ranks found in the historical material, medium length, deferred recall.

5th. The high school group showed the highest coefficient of correlation between the general scholastic ranks and the ranks in the descriptive material, medium length, immediate recall.

6th. The college group showed the highest coefficient of correlation between the general scholastic ranks and the ranks in the narrative material, medium length, immediate recall.

7th. The memory correlations are highest in the case of the elementary school pupils.

8th. The averages indicate that there is little difference between immediate and deferred memory, in this type of test, as a measure of scholastic ability.

9th. The averages indicate that a medium length passage as a test of logical memory is the best measure of scholastic ability except for the elementary pupils.

10th. The averages also indicate that the descriptive type of prose passages correlate best with the general scholastic standing. The narrative is second, and the historical third.

11th. The coefficients of correlation secured by comparing the results of the tests in logical memory with the general scholastic ranks, indicate that a logical memory test may be considered a fair measure of the general scholastic standing of the individual student.

12th. There seems to be no evidence of a practice improvement in any of the groups.

THE MEASUREMENT OF PROGRESS IN LANGUAGE ABILITY

RUDOLF PINTNER

Ohio State University

The Trabue Language Completion Tests*, Scales B and C, were used to obtain a measure of the increase in language ability of school children. Two schools were tested; the one a junior high school with an enrollment of about 750, and the other an ordinary grade school with an enrollment between 550 and 650. The junior high school would be considered a rather superior school, whereas the other would be ranked as an average city school.

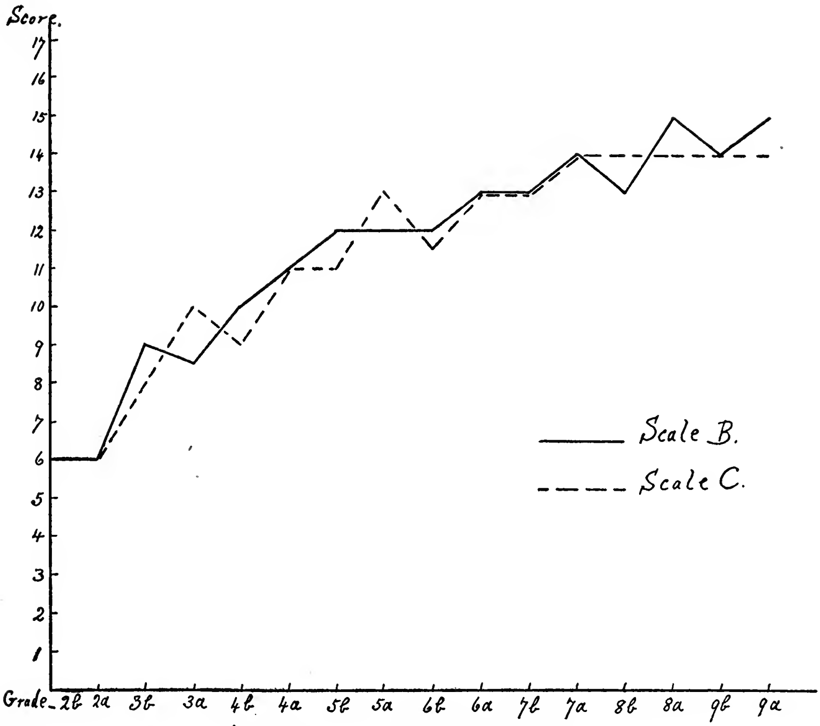
Scale B was given to the children in October, 1916, and Scale C in May, 1917. No change in the method of language instruction was made by the teachers and for the most part they were not aware of the purpose of the tests, because several other tests were given at the same time. At the time of the first test they did not know that the pupils were to be re-tested at a later date.

The procedure in giving and scoring the tests followed the standard method prescribed by Trabue.

The grade medians for each school are shown in Graphs I and II.

These medians are for the grades irrespective of the children in those grades, that is to say, the median for the 2a grade is not based upon the same children for Scale B as for Scale C, since the majority of the pupils in 2a have advanced to 3b in May when Scale C was given. The graphs show the language ability for each grade in the school and unless any radical change in language teaching has taken place, we would not expect to see any decided difference between the two scales. Graph I shows the results for the junior high school and Graph II the same for the grade school. According to expectation we do not see any decided difference between the curves for the two scales. The fluctuations are such as might arise owing to the differing abilities of the children making up the separate classes. The two schools are shown on separate graphs because the overlapping of the four curves on one graph would be too confusing. A comparison of the two graphs, however, will show that on the whole the grade school (Graph II) is slightly lower than the junior high school (Graph I).

*Trabue M. R. *Completion Test Language Scales.* Teachers College Contributions to Education, No. 77, New York, 1916.



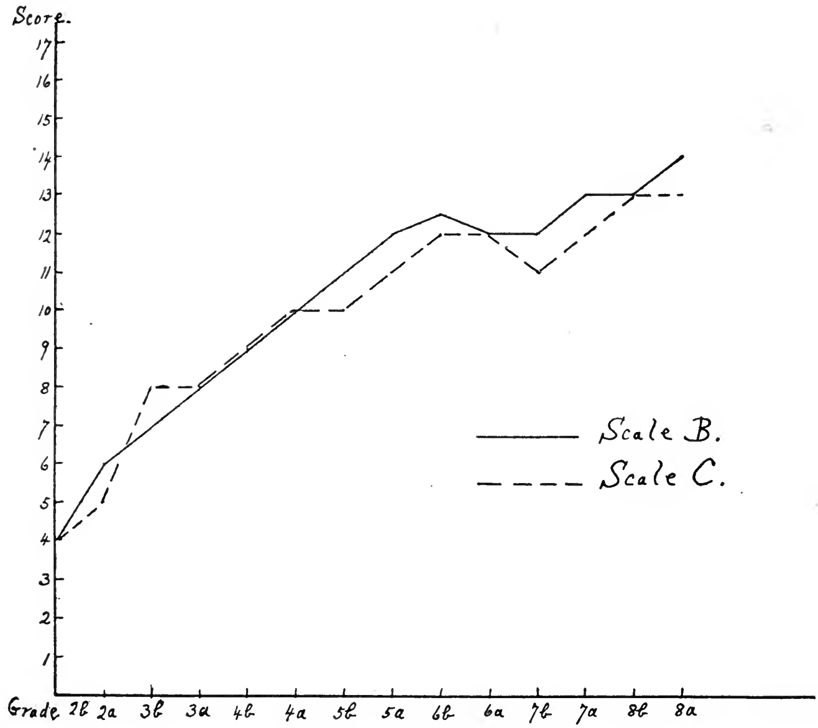
GRAPH I. JUNIOR HIGH SCHOOL

Table I shows the medians for each grade, irrespective of the

TABLE I.

Grade Medians for Scales B and C

	Junior High School		Grade School	
	B.	C.	B.	C.
Grade	October	May	October	May
2	6	6	5	4
3	8.75	9	7.5	8
4	10.5	10	9.5	9.5
5	12	12	11	11
6	12.5	12.25	12.25	12
7	13.5	13.5	12.5	11.5
8	14	14	13.5	13
9	14.5	14		
Number of Cases...........	751	746	558	634



GRAPH II. GRADE SCHOOL

different sections of each grade. These medians give us the normal grade ability in language for the two schools. In both cases there is very little difference between Scales B and C, showing that the scales are roughly equal in difficulty. Only in two cases is there a difference of one point, namely in the second and seventh grades in the grade school, where one point less was scored on Scale C than on Scale B. This may result either from the greater difficulty of Scale C, mentioned by Trabue, or from a poorer group of children in those grades in May as compared with October. On the whole, however, the table seems to show that a school keeps a certain standard of language ability from year to year.

A first rough comparison of relatively the same group of children can be obtained from comparing each grade on the B Scale with the next higher grade on the C Scale, because the children as a whole have progressed from the one grade to the higher grade during the period covered by the tests. This comparison is shown in Table II. The table is to be interpreted as follows: In the junior

TABLE II.

Grade Medians for B and C Compared

	Junior High School			Grade School		
Grade	B.	C.	Difference	B.	C.	Difference
2b–2a	6	6	0	4	5	+1
2a–3b	6	8	+2	6	8	+2
3b–3a	9	10	+1	7	8	+1
3a–4b	8.5	9	+0.5	8	9	+1
4b–4a	10	11	+1	9	10	+1
4a–5b	11	11	0	10	10	0
5b–5a	12	13	+1	11	11	0
5a–6b	12	11.5	–0.5	11	11	0
6b–6a	12	13	+1	12	12	0
6a–7b	13	13	0	12.5	12	–0.5
7b–7a	13	14	+1	12	11	–1
7a–8b	14	14	0	12	12	0
8b–8a	13	14	+1	13	13	0
8a–9b	15	14	–1	13	13	0
9b–9a	14	14	0			
			Ave. = +0.47			Ave. = +0.32

high school the 2b grade in October became the 2a grade in May, and made a median score of 6 on Scale B in October and a score of 6 on Scale C in May, showing a difference of zero, while in the grade school the same grade increased from 4 to 5 showing a difference of plus one; the next group of children (2a to 3b) showed the same medians and the same increase in both schools. It will be noted that on the whole the lower grades show some increase in the median, whereas the upper grades show hardly any. The steady increase in the median seems to stop at the fourth grade. Above this the greater number of grades show a difference of zero. The average of all the differences shown in columns four and seven shows a gain of +0.47 for the junior high school and +0.32 for the grade school.

Eliminating now all individuals who did not take both tests, we find that we have 598 in the junior high school and 418 in the grade school. Tabulating the individual gains or losses in score from Scale B to Scale C, and averaging these gains or losses for each grade, we find the results shown in Table III. In the junior high school in the group of 36 children that moved from 2b to 2a the average gain in score was +2.31 points and the median gain +2; in the corresponding group in the grade school the gain was +0.81 and the median gain zero, and so on for the rest of the table. Again we note that there is a constant and steady gain in the lower classes up to and including grade 4, but from there on the change from Scale

TABLE III.

Individual Gains or Losses

Grade	Junior High School			Grade School		
	Average	Median	No. Cases	Average	Median	No. Cases
2b—2a	+2.31	+2	36	+0.81	0	21
2a—3b	+2.26	+2	39	+1.50	+2	34
3b—3a	+1.79	+2	38	+0.53	+1	28
3a—4b	+0.67	+1	27	+0.61	0	28
4b—4a	+2.00	+2	40	+0.64	0	33
4a—5b	−0.25	0	24	−0.10	0	40
5b—5a	+1.06	+1	34	−0.24	0	51
5a—6b	−0.79	−1	28	+0.36	0	33
6b—6a	+0.29	+0.5	28	−0.06	0	34
6a—7b	−0.03	0	33	+0.18	−1	22
7b—7a	+1.29	+1	48	−0.38	−0.5	2[illegible]
7a—8b	+0.08	0	36	−0.77	−1	35
8b—8a	+0.87	0	63	−0.36	0	33
8a—9b	−0.97	−1	33			
9b and a mixed	−0.73	−1	41			
9b—9a	+0.08	+0.5	50			
		Total	598			418

B to Scale C fluctuates around zero, sometimes being positive and sometimes negative. Again we note that the junior high school is superior to the grade school in amount of progress made.

The distribution of all the cases, irrespective of grade, according to the number of points gained or lost is shown in Table IV. The table shows that in the junior high school two individuals, or 0.3 per cent. of the total, made a gain of 9 points (*i. e.* the score on Scale C was 9 or more than the score on Scale B); and, further, third line, that 4, or 0.7 per cent. of the junior high cases, made a gain of 7 points, and one of the grade school children made this gain, that is, 5 of all the cases under consideration, or 0.5 per cent., made a gain of 7 points. The rest of the table is to be interpreted in the same way. A comparison of the two schools is well brought out on Graph III. The junior high shows a greater number of cases making some kind of a gain and the grade school shows a greater number at zero and below. The combined curve for the two schools would, of course, run between the two curves on the graph and would show a very normal distribution.

From Table IV we can compute the average and median gain or loss for all the children for the two schools separately and combined:

TABLE IV

Distribution According to Points Gained or Lost

Gain or Loss	Junior High No.	Junior High Percent.	Grade School No.	Grade School Percent.	Both Schools No.	Both Schools Percent.
+9	2	0.3			2	0.2
+8			2	0.5	2	0.2
+7	4	0.7	1	0.2	5	0.5
+6	8	1.3	9	2.1	17	1.6
+5	21	3.5	10	2.4	31	3.0
+4	50	8.4	16	3.8	66	6.5
+3	63	10.5	36	8.6	99	9.8
+2	84	14.1	45	10.8	129	12.7
+1	86	14.4	55	13.2	141	13.9
0	91	15.2	78	18.7	169	16.7
−1	68	11.4	56	13.4	124	12.2
−2	54	9.0	56	13.4	110	10.9
−3	37	6.2	25	6.0	62	6.1
−4	16	2.7	21	5.0	37	3.6
−5	6	1.0	1	0.2	7	0.6
−6	1	0.2	7	1.7	8	0.7
−7	4	0.7			4	0.4
−8	3	0.5			3	0.3
Totals	598	100.1	418	100.0	1016	99.9

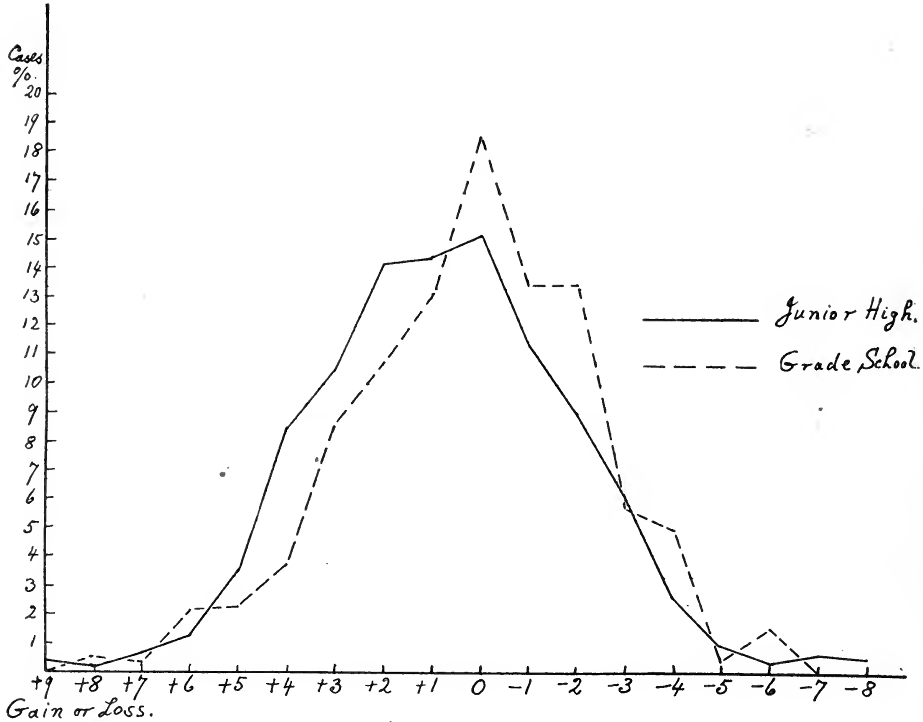
	Average	Median
Junior High School..........	+0.67	+1
Grade School................	+0.17	0
Both Schools................	+0.46	0

The individual gain is so slight that it does not show itself in the median of all the cases together. It is on the average half of one point.

We have noticed, however, all along that the marked increase was in the lower grades and that the loss in ability as measured by the scale occured in the upper grades. We have, therefore, computed the gain in score for the cases from the second through the fourth grade (October standing). This gives us the following results:

	Average Gain
Junior High School..................	+1.87
Grade School........................	+0.92
Both Schools........................	+1 48

With these younger children we have a gain of about one and one-half points. In the better class school it is about two points and in the other school about one. It is obvious, therefore, that the



GRAPH III. INDIVIDUAL GAINS AND LOSSES

scale is too easy for the higher grades and, therefore, does not allow of a measurement of their progress in language ability. Furthermore, the progress in language ability is very probably not so rapid as the child grows older. Since the sentences in Scales B and C are more or less so constructed as to show a difference of two points from grade to grade, we may say that in these lower grades this rate of improvement is practically taking place. A gain of 1 5 has been made from October to May leaving the additional increase of 0.5 to be made during the rest of the year, when, perhaps, language progress may not be so rapid owing to the smaller amount of formal instruction by the school during the summer months.

Finally we give below the grade medians for Scale C for all the pupils tested for purposes of comparison with other work. The grade medians for Scale B are not published here, since the writer has published elsewhere* grade medians for this scale for a much larger group of cases. The grade medians for Scale C are:

*PINTNER, R. *The Mental Survey.* D. Appleton and Company, 1918.

Grade	II	III	IV	V	VI	VII	VIII	IX	
Median	4	8	10	11	12	13	13	14	
No. of Cases	130	180	161	204	148	170	242	145	1380

Summary

1. Two schools were tested at the beginning and end of the school year by means of the Trabue Scales B and C.

2. Differences in the language ability and in the language progress of the two schools were marked and constant.

3. In both schools children above the fifth grade showed no language improvement as measured by the scales. It would seem that Scales B and C are not suitable for the measurement of language progress of older pupils during a relatively short interval of time. Harder scales may bring out better what progress may exist.

4. The language progress of the total group was about half a point.

5. The language progress of the younger children in grades four and below was 1.5 points, about what is to be expected in measuring the development for about seven months of the school year.

TESTS OF APPLICANTS FOR ADMISSION TO UNIVERSITY OF MINNESOTA MEDICAL SCHOOL

M. E. HAGGERTY

University of Minnesota

Candidates for entrance to the University of Minnesota Medical School in September, 1916, were given two tests with a view to determining mental capacity and general mental fitness to pursue the work of the school. The tests were given not so much for the purpose of accepting or excluding the immediate applicants as to provide a method and standards for the proper evaluation of future candidates. Since similar tests are being repeated with students applying for admission in September, 1917, the present paper should be regarded as a report of an investigation in progress.

No applicants were examined who did not meet the academic requirements for entrance to the Medical School. These requirements are substantially as follows: A high school course including two years of Latin, two years of prescribed and elective studies in the College of Science, Literature and the Arts in the University of Minnesota or its equivalent. During these two years of work the following studies are prescribed: rhetoric, 6 credits; physics, 8 credits; general chemistry and qualitative analysis or these subjects combined with organic chemistry, 12 credits; zoology, 6 credits; and a reading knowledge of German.

The Tests

Two types of tests were given. One was an omnibus test fashioned after the one reported by Scott.* This test was arranged in triplicate since the students were tested on different days and it was desired, so far as possible, to prevent communication from one group to another.

As a second test, sets 9, 9⅓, and 10⅓ of the Thorndike† Improved Scale for Reading Ability, Scale Alpha 2, were used.

Methods of Testing

The candidates, 93 in all, were tested either individually or in groups ranging up to 25. Both the omnibus test and reading test

*Scott, W. D., *The Scientific Selection of Salesmen.* Advertising and Selling Magazine, 1916.

†Thorndike, E. L., *An Improved Scale for Measuring Ability in Reading. Teachers' College Record,* November, 1915, and January, 1916.

were given at one sitting. The former was given first and the paper was handed in before the reading test was given out. Accurate time was kept on both tests.

Scoring

The tests were scored both for rate and quality. In the omnibus tests each error was penalized by adding ten seconds to the raw time, thus giving a "weighted time." The three omnibus tests were rated as of the same difficulty. Subsequent studies with the same tests on approximately 2000 high school students show that there was no substantial error in so doing. In the reading tests the subjects were credited with the values achieved on the Thorndike scale as well as being scored for time. The Thorndike values are called "quality of reading" in the discussion to follow.

Results

At the time the tests were given there were no data available to show what high or low scores in either of the tests signify in terms of probable success in the work of a medical curriculum. Common sense indicates that in a course as crowded as that of the medical school the quality and rate achieved in difficult reading would be an important indication of probable success in that course. Certain experimental evidence indicates that the rate of ordinarily difficult associations is some indication of general intelligence. That the two tests measure in part the same functions is indicated by the correlation between them of .34 (weighted time in association and quality of reading. (See Table I).

This same correlation figure indicates that in part the two tests measure different things. It was thought, therefore, that a high rank in the two tests would be a double indication of superior ability and would clearly indicate the acceptability of the applicant. Low rank in both tests would argue for rejection on the basis of mental incompetence. Exception to this general rule was made in the case of two students, one Russian and one Hindu, whose language difficulties were inordinate. As a matter of fact no students were rejected wholly on the basis of the tests, as there was always "corroborative evidence" of low ability from the academic records and general information tests.

Of the 93 persons taking the tests, 69 were accepted for entrance and actually entered the medical school in September, 1917. Concerning these we have now available the records for the work of the first semester in the medical school. These records together with the pre-medical record in the academic college and the two tests are the basis of the following table of correlations.

TABLE I.

	S. L. A. Marks	Medical School Marks	Medical School Marks, 15 hrs. cre.	Medical School Marks, 14 hrs. cre.	Quality of Reading	Time of Reading	Omnsbus Test Raw Time	Omnibus Test Weighted Time	Weighted Time and Quality of Reading	Weighted Time and Time of Reading	S. L. A. and Medical School Marks	Omnibus Test, Errors
S. L. A. Marks........		49	59	51	46	—14	39	35	55	36	85	36
Medical School Marks	49				62	24	51	60	65	57	85	29
Medical School Marks, 15 hrs. cre.........	59				77	24	68	53	81	61	88	17
Medical School Marks, 14 hrs. cre.........	51				56	4	46	47	69	43	85	45
Quality of Reading...	46	62	77	56		21	41	34	82	41	55	36
Time of Reading.....	—14	24	24	4	21		19	10	17	76	4	—18
Omnibus Test, Raw Time.............	39	51	68	46	41	19						39
Omnibus Test Weighted Time...........	35	60	53	47	34	10						67
Weighted Time and Quality of Reading.	55	65	81	69	82	17						51
Weighted Time and Time of Reading...	36	57	61	43	41	76						33
S. L. A. and Medical School Marks......	85	85	88	85	55	4						25
Omnibus Test Errors.	36	29	17	45	36	—18	39	67	51	33	25	

For a proper understanding of this table some explanation of terms is necessary. "S. L. A. marks" means the quality of work done by the student in his college course prior to his entrance to the medical school. In the University of Minnesota the grades which a student receives carry what are called "honor points." Grades of A, B, and C merit three, two, and one honor points respectively. A grade of D or below carries no honor points. A student must have an average of one honor point per credit hour to receive a degree from the University. In computing the "S. L. A. marks" in this study the total number of honor points made by a student was divided by his number of credits. These marks therefore indicate the average quality of work done by the student in his pre-medical course. In no case was the amount of such academic work less than two full years. From this it ranged up to four years, credited with the bachelor's degree. Sixty-eight cases were available for correlation.

"Medical marks" are the marks received by the student during the first semester in the medical school. Only students carrying full work, 14 or 15 credit hours, were considered. The difference in number of credits was determined by the fact that some students took organic chemistry, a four-hour course, and others took physiological chemistry, a three hour course. There were a considerable number of irregular students; some had secured medical school credits during their pre-medical courses; others did not carry full work in the medical school during the semester in question All such students were excluded in figuring correlations. The number available for correlations were 24 with fifteen credits and 26 with fourteen credits or 50 in all.

In figuring the correlations it was desirable to credit quality as well as quantity of work. Since a very large number of students were reported with marks of D which carried no honor points at all it was clearly unfair to use the same method of figuring quality of work as was used on the "S. L. A. marks." Accordingly the following arbitrary values were assigned to the several letters:

A=10 B=8 C=5 D=2 F=0

The "medical school marks" were computed by finding the number of such points to which a student was entitled and dividing the number by the number of credits. Thus: 3 credits A=30 points, 5 credits B=40 points, 6 credits D=12 points, a total of 82 points. This total divided by 14, the number of credits, equals 58.5, the actual mark of quality to which the student is entitled. By this method the 15 credit and 14 credit students may be considered together without serious statistical error.

"Quality of reading" means score on Thorndike scale. "Time of reading" means total time required for reading test. "Omnibus Test: Raw time" equals time required for the association test. "Omnibus Test: Weighted time" is the same with 10 seconds added for each error. Available for correlation in all of these measures were 68 cases.

It was originally thought that the weighted time in the omnibus test and the quality of reading would be the best prognostic measures, and the applicants were rated in them first. As it turns out the correlation of these combined measures with the marks subsequently made by these students is .65 for the group of 50, .81 for the fifteen-credit cases and .69 for the 14 credit cases. Either of these tests alone gives a slightly lower but still substantial correlation. In no case is it less than .56.

A comparison of the correlations of S. L. A. marks and medical school marks shows that the tests were a much better indication of ability to do the work of the medical school than were the academic records of these applicants. There is not a single case in which the test correlations are so low as the academic mark correlations. In fact either test alone is a better prognostic measure than were the S. L. A. marks. The "quality of reading" correlates with medical marks to the extent of .62, the omnibus test to .60 and the two combined to .65, while the S. L. A. marks show but .49.

The prognostic value of the tests may be seen also from the per cent of median, tertile and quartile retention which turns out to be 71, 56 and 46 respectively as shown in Figures 1, 2, and 3. Thus in Figure 1, 23 persons are shown to stand in the upper half of the group in both of the tests and in the upper half of the group in medical school marks; 16 persons stood in the lower half of the group in both the tests and medical school marks. Nine persons were in the first half in the tests but in the latter half in school marks, while 2 were in the lower half in the tests but in the upper half in school marks. From Figure 3 it is apparent that no person who fell to the second or third quartile in the tests achieved a place in the first quartile in school marks. Only two in the second quartile of the tests came into the school mark first quartile. On the other hand no person who achieved a first quartile in the tests fell to the fourth quartile in school marks. But two of this test group went to the third quartile while 6 are found in the second and 10 in the first quartile. Only 9 of the 50 individuals are displaced in the marks more than one quartile from their test rankings. Such displacement as does exist is largely due to the very great number of D's given in the anatomy courses.

As in the case of the correlations, the median, tertile and quartile retentions show the tests, either singly or combined, to have been better prognostic instruments than the academic records of the students in the Arts college.

FIGURE 1.

Median Retention

RANKS IN MEDICAL SCHOOL MARKS

Halves

Standing in tests. Weighted time plus quality of reading.	1	2
1	23	9
2	2	16

Median retention 71%

These results are significant. They mean that an entrance examination board can determine by three hours work the fitness of one hundred applicants for the work in a medical school more exactly than they can derive such information from the laborious examination of academic records, often hard to obtain and equivocal in meaning.*

A curious fact is the somewhat equivocal correlations shown by the "time of reading" scores with academic marks. With the S. L. A. marks it is -14 and with medical school marks it is positive .24. From this it would seem that in a test as difficult as the reading scale the time consumed is no indication of ability. This was not true of the association test where the "raw time" scores gave positive correlations of .39 and .51 with S. L. A. and medical school marks, .68 and .46 with the 15 and 14 credit student's marks and .41 with the quality of reading. This probably means that in the simpler associations the rate is more important than in the more difficult tests and more important than the quality of work as measured by the errors in this test.

It is probably fair to interpret all these results as follows: As a means of prognosis for success in the medical school course, it is

*The giving of the two tets occupied the time of two examiners for not to exceed one and one-half hours. Since the applicants may be tested in groups, a hundred may be examined at a sitting. The scoring of the tests may be done very rapidly by persons familiar with the work.

FIGURE 2.

Tertile Retention

RANKS IN MEDICAL SCHOOL MARKS

Tertiles

Standing in Tests. Weighted time plus quality of reading	1	2	3
1	13	6	6
2	2	7	5
3	0	5	8

Tertile Retention 56%

FIGURE 3.

Quartile Retention

RANKS IN MEDICAL SCHOOL MARKS

Quartiles

Standing in tests. Weighted time plus time of reading	1	2	3	4
1	10	6	2	0
2	2	5	1	6
3	0	1	3	4
4	0	1	4	5

Quartile Retention 46%

important to know the rate and quality of fairly simple associative processes (weighted time in omnibus test) and also to know how difficult an associative problem a student can master more or less regardless of the time required to do it (quality of reading).

The desirable amount of correlation between mental tests and academic achievement is difficult to fix. Success in school work is apparently dependent upon many factors, such as good observational powers, memory, imagination and intelligence or reasoning capacity. Presumably such tests as the ones here used are directed to the measurement of such intellectual functions. Beyond these, however, are other mental functions such as industry, endurance, interest, persistence of motive, freedom from prejudice, physical fitness, general tonicity of muscular system, and a wide range of instinctive and emotional states and activities illy defined but fairly described as "attitudes," "sets," etc. It these tests measure merely the intellectual functions, then a certain portion of the student's mark is due to mental abilities not evaluated by the tests. In such cases the correlation should fall below 1, or a perfect figure. The amount it falls short would be determined by the degree to which such non-intellectual functions determine the school mark. In such cases, a correlation figure of .60, .70 or .80 might be as high as could be expected and should be regarded as the maximum to be expected from tests of this type. In case .80 were determined as such maximum, a correlation of .60 with a test for intellectual functions would be explained as due to the inadequacy of the test for the measurement of such functions. In such case it would be necessary to improve the test or to supplement it with other tests for intellectual capacities. With the fixed maximum of .80 a test showing .80 would be regarded as serving all the legitimate ends of the intellectual test. Its proper supplement would be tests for the instinctive, emotional and other factors to which a certain portion of the student's mark is due.

It is extremely doubtful, however, whether such vivisection of mental capacities is the way to their proper evaluation. Association in an individual is not the abstract thing we talk about in our text books but is shot through with the very interests and emotions we so carefully discriminate it from in our discussions. When a person attacks the "hard directions" test it is not merely the abstract processes of association, memory, discrimination, and attention which are involved. It is all of these modified by the desire to suc-

ceed, to do the task quickly, to obviate errors, and suffused with the affective and emotional states prompted by the surprise, absurdity and complexity of the test. The time score recorded at the end of the test is, therefore, a measure of the complex mental state involving all of these functions. Similarly, when an individual works diligently for 30 minutes on the Thorndike reading scale it is not fair to assume that the result of his work is merely an intellectual product. It is a mental product measuring his intellectual capacities, to be sure, but also indirectly testing his interest, his endurance, persistence and other emotional and instinctive attitudes. The implication of those considerations is that a correlation of .70 between a test and school marks means that the test measures more than the intellectual factors operating to produce the marks but measures them all partially. It means further the possibility of a test which will correlate much higher with actual performance than would be the case if the test were of purely intellectual functions. Conceivably the correlation could approach a perfect relationship. This means that none of the measures available from these tests is a satisfactory one. Valuable as they are it is probable that they may be improved so as to yield even higher correlations than the ones here reported. Two changes, on the basis of experience, seem worth while. The reading test should be presented with more degrees of difficulty so as to distribute the individuals more widely. The association test should be made of more uniform difficulty throughout. This means that some substitution should be made for the first three parts which are apparently so easy for students of the advancement of those here considered that the mechanical performance of writing the answers consumes more time than the mental processes of thinking the correct response.

In the testing of the students applying for entrance to the Medical School in September 1917 these changes have been made. The value of these changes is yet to be determined.

TESTING PRACTICE MATERIAL IN THE FUNDAMENTALS

CYRUS D. MEAD
College for Teachers, University of Cincinnati
and
CHARLES W. JOHNSON
Williams Avenue School, Norwood

One of the contributions which the science of measurements and tests makes to the field of education is the tools it offers for evaluating different methods of instruction (or learning) and different materials which enter into the teaching or learning process. Not many years ago, the only means of determining merit in material or method was by judgment or opinion, based, perhaps, upon so-called experience or formal examination. Such a judgment might vary as the convictions or prejudices of the teacher or administrator varied. The construction and application of standardized tests puts a somewhat impersonal instrument in the hands of the teachers so that the results obtained by the use of such tests are susceptible of less variable interpretation than the more familiar examination. The challenge put to any claimant in educational matters nowadays is, upon what data or results do you base your statements? The more impersonal such data are, the more weight such statements carry.

The purpose of this experiment was to obtain results upon which an opinion could be based as to the comparative degree pupils were made proficient by the use of two kinds of practice material, not only proficient in the fundamentals of arithmetic, but in the so-called arithmetical reasoning phases. The latter problem will be taken up in the second part of this study. Such an experiment as regards proficiency in the fundamentals had previously been made with nearly one thousand fifth grade pupils in the Cincinnati schools working in twenty-four classes of fourteen schools. In a few months' time, using not more than fifteen minutes daily (which included distributing and collecting of material, and recording scores), most of these pupils were advanced beyond sixth grade standards regardless of the kind of material used. The experiment complete, with tabular results, has been issued as a *School Efficiency Monograph* by the World Book Company. This large group experiment probably allowed varying factors to enter into the results, although the

conditions were controlled as nearly as possible. The direct purpose of this test then was to carry out the same experiment with smaller groups under more direct supervision and control. This problem was assigned, accordingly, to three members of a graduate seminar in Education,* each man being the principal of a large elementary school. To them chiefly is due the credit for a very painstaking and thorough piece of work.

Six classes were selected, four fifth grades and two sixth grades. They were paired so that three used one kind of practice exercises and three another kind. The material used was the Courtis Standard Practice Tests (Exercises), and the Thompson Minimum Essentials. Each principal supervised two classes, one class having one kind of material, the other, the second kind. The experiment extended over a period of about ninety school days,beginning about the first of December, 1916, and closing near the middle of April, 1917. In order to measure progress made by two groups, an identical test must be given to each group preliminary to and directly after the different practice materials are used. By comparing the median gains (or losses), or the percentile growths, one should be able to form a fairly accurate estimate of the changes due to each of the kinds of practice material used between the preliminary and final tests. The initial ability of the six classes in the fundamentals was obtained by giving the Courtis Standard Research Series B test. (See Table IV and Charts I and II.)

On the first Monday in December each class began its drill work with the kind of material assigned. Both groups of classes devoted approximately the same time to this formal drill, ten minutes daily, such time being taken from the regular sixty minute daily period. The drill was always given at the beginning of the period. However, the pupils were allowed to use their practice sheets outside of class, if they wished, the slower ones in fact were even urged to do so. This is a perfectly legitimate use of practice material and is supposedly one of its merits. If such material has in it the elements of interest sufficient to stimulate pupils to a more diligent use, then just so much more valuable does such material become. In April the final Series B tests were given. The scores are given in the following tables:

*Charles W. Johnson, Williams Avenue School, Norwood; H. L. Crane, Oakley School, Cincinnati; D. S. Richards, Linwood School, Cincinnati.

TABLE I

Frequency Tables—Tests I and II (Series B)

OAKLEY SCHOOL — ATTEMPTS—TWO 5TH GRADES

Score	Courtis' Practice Pads								Thompson's Essentials							
	Addition		Subtrac.		Multip.		Division		Addition		Subtrac.		Multip.		Division	
	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
0	..	..	..	..	1	..	1	2	..	..	..	..	..	..	..	..
1	..	..	..	..	0	..	2	1	..	..	..	..	..	..	2	..
2	3	1	..	1	3	1	5	4	..	..	..	..	..	..	1	7
3	3	1	1	0	3	3	9	6	2	..	..	1	3	6	6	11
4	3	3	4	0	5	4	11	10	4	1	1	1	3	4	16	8
5	10	4	2	1	8	6	5	7	2	2	6	2	9	4	5	5
6	8	5	13	6	11	9	2	9	11	3	9	10	11	11	4	3
7	5	4	8	5	4	9	4	2	8	3	11	7	8	11	1	3
8	5	3	7	7	5	4	1	2	2	8	5	9	1	0	0	..
9	5	6	4	6	2	4	1	..	3	4	2	4	0	0	0	..
10	0	5	2	5	1	0	1	..	1	4	0	0	1	1	1	..
11	1	1	1	5	..	3	1	..	2	3	2	2	0	.	1	..
12	..	2	1	2	..	..	..	..	1	2	1	1	1	..	..	..
13	..	3	..	1	..	..	..	..	1	4	..	..	..	..	..	..
14	..	3	..	0	..	..	..	..	..	0	..	..	..	..	..	..
15	..	1	..	1	..	..	..	..	..	2	..	..	..	..	..	..
16	..	0	..	1	..	..	..	..	..	1	..	..	..	..	..	..
17	..	1	..	1	..	..	..	..	..	..	..	..	..	..	..	..
18	..	..	..	1	..	..	..	..	..	..	..	..	..	..	..	..
19	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
20	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	43	43	43	43	43	43	43	43	37	37	37	37	37	37	37	37
Median	6.4	9.2	7.3	9.3	6.2	6.9	4.5	4.9	7.0	9.5	7.2	7.7	6.3	6.5	4.6	4.1
Gain		2.8		2.0		.7		.4		2.5		.5		.2		—.5

OAKLEY SCHOOL — ACCURACY—TWO 5TH GRADES

Score	Courtis' Practice Pads								Thompson's Essentials							
	Addition		Subtrac.		Multip.		Division		Addition		Subtrac.		Multip.		Division	
	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
100	3	2	5	5	3	6	10	10	3	3	10	6	7	4	10	6
90	0	0	1	1	0	0	0	0	0	2	0	0	0	0	1	0
80	4	4	11	8	8	4	5	5	7	6	11	5	14	4	3	2
70	2	10	5	9	9	4	7	5	2	3	3	6	2	5	6	2
60	5	6	6	4	5	3	7	3	4	5	4	4	4	8	3	6
50	5	9	6	7	5	7	3	6	9	5	4	8	5	6	2	5
0—49	24	12	9	9	13	19	11	14	12	13	5	8	5	10	12	16
	43	43	43	43	43	43	43	43	37	37	37	37	37	37	37	37
Median	46	62	72	72	68	54	71	67	58	62	83	68	82	64	73	56
Gain		16		0		—14		—4		4		—15		—18		—17

TABLE II

Frequency Tables—Tests I and II (Series B)

LINWOOD-WILLIAMS AVENUE SCHOOLS — ATTEMPTS—TWO 5TH GRADES

Score	Courtis' Practice Pads								Thompson's Essentials							
	Addition		Subtrac.		Multip.		Division		Addition		Subtrac.		Multip.		Division	
	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
0	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..
1	..	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..
2	1	..	..	..	3	1	6	1	1	..	..	..	..	..	1	1
3	3	1	1	..	6	1	9	1	2	1	..	..	..	..	3	2
4	6	0	6	..	9	0	11	6	2	1	1	..	3	3	3	3
5	9	0	5	..	10	4	3	5	7	2	2	..	5	3	5	4
6	4	3	5	1	3	7	3	9	6	7	6	4	5	5	1	2
7	4	6	9	7	3	4	..	2	2	4	2	2	6	4	5	2
8	6	2	7	5	..	6	..	1	3	3	5	7	2	1	2	5
9	0	3	0	1	..	5	..	1	4	3	3	4	1	7	2	0
10	0	2	0	8	..	2	..	0	1	3	2	3	5	2	3	4
11	1	1	0	5	..	1	..	1	..	2	1	1	0	1	1	0
12	..	5	1	4	..	..	..	2	..	0	2	2	0	1	0	0
13	..	1	..	..	..	..	..	2	..	2	3	0	0	0	1	3
14	..	4	..	..	..	..	..	..	..	..	0	1	0	0	0	0
15	..	0	..	..	..	..	..	..	..	..	0	3	00	0	0	0
16	..	2	..	..	..	..	..	..	..	..	0	0	0	0	0	0
17	..	1	..	..	..	..	..	..	..	..	1	1	0	0	1	1
18	..	..	..	..	..	..	..	..	..	..	..	..	1	1	..	0
19	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1
20	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	34	31	34	31	34	31	34	31	28	28	28	28	28	28	28	28
Median	5.8	10.5	7.0	10.3	4.9	7.8	4.0	6.3	6.3	7.8	8.6	9.3	7.2	7.8	7.2	8.0
Gain		4.7		3.3		2.9		2.3		1.5		.7		.6		.8

LINWOOD-WILLIAMS AVENUE SCHOOLS — ACCURACY—TWO 5TH GRADES

Score	Courtis' Practice Pads								Thompson's Essentials							
	Addition		Subtrac.		Multip.		Division		Addition		Subtrac.		Multip.		Division	
	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
100	4	5	5	4	1	4	5	7	2	3	7	11	4	10	7	9
90	0	2	0	5	0	1	0	1	0	1	2	3	2	1	4	2
80	6	8	2	9	1	4	3	11	3	1	10	8	10	10	7	5
70	3	5	4	5	5	6	2	1	2	6	4	4	5	4	2	3
60	2	2	4	5	5	7	3	1	3	4	1	1	3	3	1	1
50	4	6	4	2	3	5	4	6	5	4	3	1	3	0	1	2
0—49	15	3	15	1	19	4	17	4	13	9	1	0	1	0	6	6
	34	31	34	31	34	31	34	31	28	28	28	28	28	28	28	28
Median	55	80	55	83	45	70	50	84	52	63	85	90	82	87	86	84
Gain		25		28		25		34		11		5		5		—2

TABLE III

Frequency Tables—Tests I and II (Series B)

WILLIAMS AVE.-LINWOOD SCHOOLS — ATTEMPTS—TWO 6TH GRADES

Score	Courtis' Practice Pads								Thompson's Essentials							
	Addition		Subtrac.		Multip.		Division		Addition		Subtrac.		Multip.		Division	
	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
0	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	1	..
2	..	..	..	..	..	..	..	..	..	..	..	..	..	1	1	1
3	..	..	..	..	..	..	..	..	..	..	..	..	3	0	8	0
4	1		..	..	..	..	1	..	4	..	1	..	3	1	6	5
5	8	..	1	..	1	..	4	..	5	1	7	..	5	0	6	3
6	3	..	1	..	4	..	1	..	10	3	8	4	12	10	6	8
7	5	..	4	..	5	..	5	1	7	5	9	6	9	7	1	2
8	4	4	6	..	4	3	6	3	8	7	7	9	3	6	3	3
9	5	6	10	..	5	5	1	2	2	5	4	4	2	0	4	6
10	4	7	1	4	4	2	10	4	2	5	1	7	1	7	1	3
11	1	3	0	4	5	7	1	5	0	5	1	3	0	3	1	3
12	..	5	6	4	2	2	0	4	1	4	1	2	1	0	1	1
13	..	2	1	5	0	4	1	4	..	2	..	0	..	2	..	2
14	..	2	1	5	1	4	1	1	..	0	..	0	..	0	..	0
15	..	1	..	4	..	2	..	2	..	1	..	0	..	2	..	1
16	..	1	..	1	..	0	..	3	..	0	..	4	..	..	..	0
17	..	..	..	3	..	0	..	0	..	0	..	..	..	..	..	0
18	..	..	..	1	..	1	..	2	..	1	..	..	..	..	..	1
19	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	..
20	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	31	31	31	31	31	31	31	31	39	39	39	39	39	39	39	39
Median	7.8	10.8	9.4	13.8	9.4	11.8	8.8	12.3	7.1	9.8	7.4	9.3	6.8	8.1	5.7	8.3
Gain		3.0		4.4		2.4		3.5		2.7		1.9		1.3		2.6

WILLIAMS AVE.-LINVOOD SCHOOLS — ACCURACY—TWO 6TH GRADES

Score	Courtis' Practice Pads								Thompson's Essentials							
	Addition		Subtrac.		Multip.		Division		Addition		Subtrac.		Multip.		Division	
	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
100	2	3	3	12	6	5	14	11	2	10	3	4	4	7	10	10
90	0	8	3	7	4	4	2	10	0	1	1	4	0	3	0	0
80	7	4	11	6	6	8	8	4	3	13	13	6	10	9	5	11
70	4	8	5	1	5	8	4	2	4	5	5	8	0	4	4	5
60	6	1	2	3	6	4	2	0	10	3	5	7	10	5	5	3
50	6	4	5	1	2	1	1	3	7	2	6	4	5	8	2	3
0—49	6	3	2	1	2	1	0	1	13	5	6	6	10	3	13	7
	31	31	31	31	31	31	31	31	39	39	39	39	39	39	39	39
Median	67	80	82	96	82	83	95	96	60	84	76	74	65	80	70	82
Gain		13		14		1		1		24		—2		15		12

TABLE IV

	Addition						Subtraction					
	Courtis			Thompson			Courtis			Thompson		
	Att.	Acc.	Eff.	Att.	Acc.	Eff.	Att.	Acc.	Eff.	Att.	Acc.	Eff.
Oakley 5th												
Test I	6.4	46	2	7.0	58	3	7.3	72	2	7.2	83	3
Test II	9.2	62	5	9.5	62	8	9.3	72	7	7.7	68	3
Change	2.8	16	3	2.5	4	5	2.0	0	5	.5	—15	0
Lin.-Wms. 5th												
Test I	5.8	55	6	6.3	52	4	7.0	55	0	8.6	85	11
Test II	10.5	80	13	7.8	63	7	10.3	83	10	9.3	90	25
Change	4.7	25	7	1.5	11	3	3.3	28	10	.7	5	14
Wms.-Lin. 6th												
Test I	7.8	67	3	7.1	60	3	9.4	82	0	7.4	76	0
Test II	10.8	80	10	9.8	84	15	13.8	96	39	9.3	74	3
Change	3.0	13	7	2.7	24	12	4.4	14	39	1.9	—2	3

	Multiplication						Division					
	Courtis			Thompson			Courtis			Thompson		
	Att.	Acc.	Eff.	Att.	Acc.	Eff.	Att.	Acc.	Eff.	Att.	Acc.	Eff.
Oakley 5th.												
Test I	6.2	68	5	6.3	82	5	4.5	71	2	4.6	73	3
Test II	6.9	54	0	6.5	64	3	4.9	67	5	4.1	56	3
Change	.7	—14	—5	.2	—18	—2	.4	—4	3	—.5	—17	0
Lin.-Wms. 5th.												
Test I	4.9	45	0	7.2	82	7	4.0	50	0	7.2	86	11
Test II	7.8	70	6	7.8	87	14	6.3	84	13	8.0	84	25
Change	2.9	25	6	.6	5	7	2.3	34	13	.8	—2	14
Wms.-Lin. 6th.												
Test I	9.4	82	10	6.8	65	0	8.8	95	32	5.7	70	15
Test II	11.8	83	13	8.1	80	3	12.3	96	36	8.3	82	20
Change	2.4	1	3	1.3	15	3	3.5	1	4	2.6	12	5

For purposes of comparison, general (standard) median scores are inserted. Such medians have been determined by Courtis and are based upon the May and June, 1915–16 scores of many thousand pupils. Bulletin 4, p. 48, Courtis Standard Research Tests, Detroit.

Grade	Addition	Subtraction		Multiplication		Division	
Sp.	Acc.	Sp.	Acc.	Sp.	Acc.	Sp.	Acc.
5—8.6	70	9.0	83	7.5	75	6.1	77
6—9.8	73	10.3	85	9.1	78	8.2	87

TABLE V

Frequency Tables—Tests I and II (Series B)

TOTAL—ALL PUPILS | ATTEMPTS—ALL PUPILS

Score	Courtis' Practice Tests								Thompson's Essentials							
	Addition		Subtrac.		Multip.		Division		Addition		Subtrac.		Multip.		Division	
	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
0	..	..	..	..	1	..	2	2	..	..	..	..	..	..	..	..
1	..	..	..	..	0	..	3	1	..	..	..	..	..	..	3	..
2	4	1	..	1	6	2	11	5	1	..	..	..	..	1	3	9
3	6	2	2	0	9	4	18	7	4	1	..	1	6	6	17	13
4	10	3	10	0	14	4	23	16	10	2	3	1	9	8	25	16
5	27	4	8	1	19	10	12	12	14	5	15	2	19	7	16	12
6	15	8	19	7	18	16	6	18	27	13	23	18	28	26	11	13
7	14	10	21	12	12	13	9	5	17	12	22	15	23	22	7	7
8	15	9	20	12	9	13	7	6	13	18	17	25	6	7	5	8
9	10	15	14	7	7	14	2	3	9	12	9	12	3	7	6	6
10	4	14	3	17	5	4	11	4	4	12	3	10	7	10	5	7
11	3	5	1	14	5	11	2	6	2	10	4	6	0	4	3	3
12	..	12	8	10	2	2	0	6	2	6	4	5	2	1	1	1
13	..	6	1	6	0	4	1	6	1	8	3	0	0	2	1	5
14	..	9	1	5	1	4	1	1	..	0	0	1	0	0	0	0
15	..	2	..	5	..	2	..	2	..	3	0	3	0	2	0	1
16	..	3	..	2	..	0	..	3	..	1	0	4	0	0	0	0
17	..	2	..	4	..	0	..	0	..	0	1	1	0	0	1	1
18	..	..	..	2	..	1	..	2	..	1	..	..	1	1	..	1
19	..	..	..	..	..	1	..	..	..	..	..	..	..	..	..	1
20	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	108	105	108	105	108	105	108	105	104	104	104	104	104	104	104	104
Median	6.5	10.1	7.7	10.8	6.3	8.3	4.9	6.6	6.9	9.1	7.5	8.6	6.6	7.2	5.3	6.2
Gain		3.6		3.1		2.0		1.7		2.2		1.1		.6		.9

TOTAL—ALL PUPILS | ACCURACY—ALL PUPILS

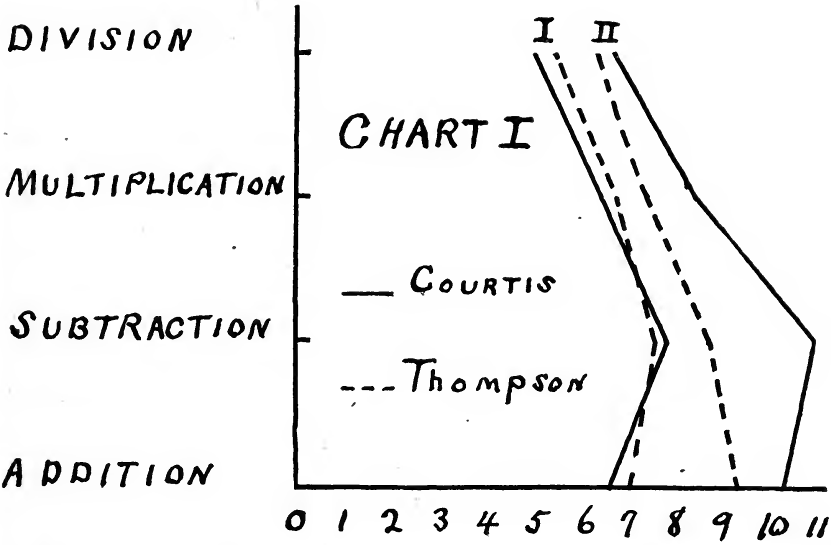
Score	Courtis' Practice Tests								Thompson's Essentials							
	Addition		Subtrac.		Multip.		Division		Addition		Subtrac.		Multip.		Division	
	I	II	I	II	I	II	I	II	I	II	I	II	I	II	I	II
100	9	10	13	21	10	15	29	28	7	16	20	21	15	21	27	25
90	0	10	4	13	4	5	2	11	0	4	3	7	2	4	5	2
80	17	16	24	23	15	16	16	20	13	20	34	19	34	23	15	18
70	9	23	14	15	19	18	13	8	8	14	12	18	7	13	12	10
60	13	9	12	12	16	14	12	4	17	12	10	12	17	16	9	10
50	15	19	15	10	10	13	8	15	21	11	13	13	13	14	5	10
0—49	45	18	26	11	34	24	28	19	38	27	12	14	16	13	31	29
	108	105	108	105	108	105	108	105	104	104	104	104	104	104	104	104
Median	56	73	71	82	66	71	75	84	57	71	81	77	79	77	76	73
Gain		17		11		5		9		14		—4		—2		—3

Table I gives the results in the fundamentals, preliminary (I) and final (II), with Series B Standard tests in two fifth grade classes of the Oakley school. It should be read: forty-three pupils of the class who practiced with the Courtis exercises made an initial median score in addition attempts of 6.4 examples, and a final median score in addition attempts of 9.2 examples, making a gain of 2.8 examples attempted. The parallel class of thirty-seven pupils who practiced with the Thompson exercises made a gain in addition attempts of 2.5 examples. The results in accuracy are read in the same manner. Tables II and III give similar data for the other paired classes. Table IV assembles the results by classes, with the score for efficiency added. Efficiency means the per cent. of pupils attempting or exceeding a certain standard number of examplesand working the same with one hundred per cent. accuracy. Some classes show a loss, as registered in the "change," in either number of attempts, or in accuracy, or efficiency.

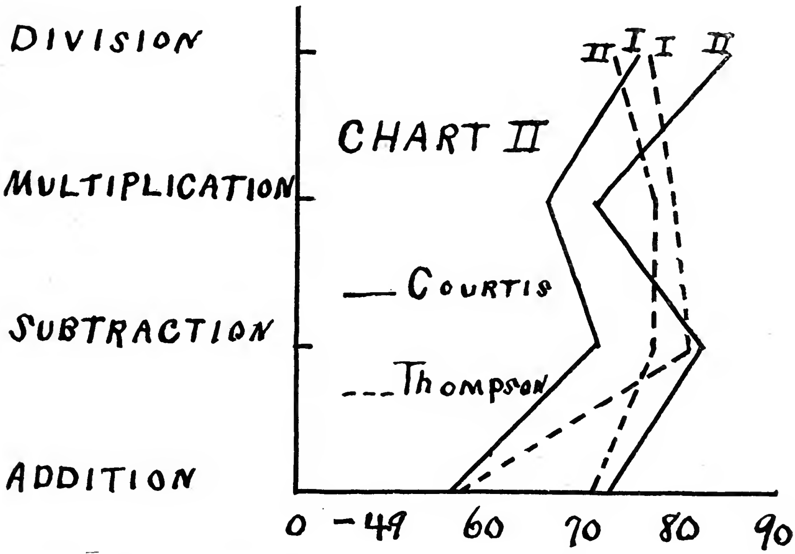
For purposes of comparison, the one hundred five pupils of the one set of exercises and the one hundred four pupils of the other set are grouped, regardless of school and grade, in Table V. These results are graphed in Charts I and II. It will be noticed that, in attempts, the pupils having the Courtis exercises made the greater increase in speed, the groups starting with about the same degree of rapidity. For accuracy of work, the Courtis group made a consistent gain in each of the fundamentals, and this based upon a greater increase in number of examples attempted. The Thompson group worked with greater accuracy at the initial test in each of the fundamentals, but, while making a consistent gain for addition, this group was less accurate at the final test in the three other processes. This shows in Chart II by the dotted lines crossing between addition and subtraction.

The Comparative Standing of the Two Groups in Initial And Final Tests in Arithmetical Reasoning

Parallel to and in conjunction with the above series of tests, data were obtained regarding any change in the ability of pupils to work so-called arithmetical reasoning problems. The Starch Scale A for arithmetical reasoning was given to all classes preliminary to and at the close of the practice drills. This scale, with tentative standards, may be found in the April, 1916, number of the *Journal of Educational Psychology.* Another set of ten problems was given



MEDIAN NUMBER OF EXAMPLES ATTEMPTED BY ALL PUPILS IN THE PRELIMINARY I AND FINAL II TESTS



MEDIAN PER CENT ACCURACY BY ALL PUPILS IN THE PRELIMINARY I AND FINAL II TESTS

also at the beginning, and for purposes of comparison, an approximately equal set weighted according to the number of pupils working, was given at the close. The question upon which light was wanted was not, does increased proficiency in the fundamentals make pupils better "reasoners" in arithmetical problems, but does it mean more accurate work when these fundamentals are applied to problems in which reasoning enters. The results in such tests are more difficult of interpretation.

Table VI gives the median preliminary and final scores for each of the classes with both the Starch Scale A and what we shall call the "combination" test, the second set of ten problems each mentioned above.

TABLE VI.

	Starch Arithmetical Scale A						Combination Test					
	C. 1	T. 1	C. 2	T. 2	C. 3	T. 3	C. 1	T. 1	C. 2	T. 2	C. 3	T. 3
Test I	5.0	7.7	4.0	7.4	11.3	7.6	8.9	9.1	11.4	18.0	56.0	34.8
Test II	7.2	9.5	7.5	10.0	13.3	10.0	27.8	41.0	33.6	46.5	72.4	54.6
Change	2.2	1.8	3.5	2.6	2.0	2.4	18.9	31.9	22.2	26.5	16.4	19.8

This table should be read: one of the classes which practiced with the Courtis material made a 5.0 median score with Scale A on the preliminary test and a median score of 7.2 with the same scale on the final test, making an increase of 2.2. The parallel class which used the Thompson material made 7.7 and 9.5 respectively, with an increase of 1.8. The results in the combination tests are read in the same manner. We have here "compared class results" similar to those given in Table IV in the tests in the fundamentals. As would be expected after five months' classroom instruction in arithmetic, there is improvement in the scores in each class. How much of this improvement was due to normal growth and how much was due to practice in the fundamentals is, of course, uncertain. However, since the "Courtis pupils" gained in accuracy in each of the four fundamentals, while the Thompson pupils lost in three of the four fundamentals (Table V), it may be that gain in accuracy in fundamentals may account for the rather favorable showing of the Courtis classes with Scale A. In the Combination reasoning tests, however, the Thompson groups appeared to advantage.

So long as median initial abilities of groups to be compared are not approximately equal, just what constitutes greater improvement is difficult to determine. Is a change from 5.0 to 7.2 better than a

change from 7.7 to 9.5? The first class has "more room" for improvement. These results with reasoning problems are given for what they might be considered worth.

Summary

One hundred five pupils practicing with one kind of practice material ten minutes daily for ninety school days, such time being taken out of the regular sixty minute arithmetic period, made a greater improvement in work in the fundamentals than did another group of one hundred four pupils under similar conditions with another kind of practice material. There was a noticeable improvement in the ability to work problems involving arithmetical reasoning in all classes regardless of practice material. The results, with regard to the fundamentals, substantiate the findings of a former experiment with nearly one thousand pupils.

The Journal of Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD.

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

NOTICE TO SUBSCRIBERS. Notify the publishers, Warwick & York, Inc., of any change of address; as copies lost in transit because of subscriber's removal to another address will not be duplicated.

THE JOURNAL is issuing a series of Monograph Supplements for the publicaiion of more extensive researches in educational psychology. For details prospective contributors may apply to H. E. Buchholz, 10 East Center St., Baltimore, Md.

The price of the JOURNAL is $3.00 per year of ten monthly issues, if paid in advance. Single issues are 50 cents each. Back volumes are $3.50 in original wrappers and $4.75 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME IX MAY, 1918 NUMBER 5

EDITORIAL

War and Credulity

While it is the proverbial last straw that is supposed to break the camel's back, most breaks occur not through the successive increments of straws, but as the result of internal stresses and strains which were unrecognized until some sudden increase in the burden revealed the weakness. An instance of this is supplied in the phenomenon of "shell shock" of which we hear so much now-a-days. The name appears to be something of a misnomer, for the condition has none of the suddenness of onset connoted by the term "shock," nor is it caused by the bursting of any particular shells. It is rather the general result of the intense strain of the firing line upon nervous organizations that are temperamentally predisposed to be affected by horrible experiences. The massing of such experiences on the susceptible individual before the organism is able to react to them by habituation produces an overpowering, deadening, stupefying effect which may result in complete mental tetanus. The large majority of the men in the firing line, if they have specific work to do, react quickly to the situation and are able to throw off the depression.

In like manner the general strain and anxiety of war reveals the latent weaknesses of the civilian community. One of the most striking of these is the greatly enhanced credulity with which even intelligent people receive the wildest and most improbable statements. One needs only to recall the stories of infected court plaster, of poisoned food, of broken glass in candy and bread, which have had such wide credence in the past year. Dwellers in Atlantic seaboard cities know with what awed fascination preposterous tales of German air raids along the American coast, of wholesale transport sinkings, of hundreds of thousands of American wounded secretly brought back to this country and crowded into local hospitals, have been eagerly imbibed and passed on to others.

A particularly sinister form of this weakness of credulity seems to be on the increase among the intellectuals of England and France in the acceptance of evidences of spiritistic manifestations. With the increasing losses of friends and relatives and the general gloom occasioned by the war, there is a quickening of religious feeling, and a readiness to lend a willing ear to stories of communication with the spirit world. This credulity is especially difficult to combat because of the impossibility of proving the negative, because of the strong emotional bias in its favor, and because of its agreement with the tenor of religious dogmas implicitly believed from childhood. No one was greatly surprised when Sir Oliver Lodge professed to have received communications from the spirit of his son, Raymond, who had been killed in the war, for Sir Oliver's belief in spiritism has long been professed. But now comes Sir Arthur Conan Doyle, backed by his prestige as a popular writer of fiction, declaring his conviction, reached since the outbreak of the war, that spiritism is the new revelation of truth to humanity, "the rebirth of the original faith and practice of Christ and his immediate followers." Likewise, we have the testimony of W. J. Crawford, professor of mechanical engineering in Queen's College, Belfast, in a book in which he claims to have applied to spiritistic phenomena "the same principles and methods of investigation and test that he would use in his laboratory upon the manifestations of any accepted and understood form of force." Finally, there has recently been translated into English a book called *The Psychology of the Future*, by the French pseudo-psychologist Emile Boirac, who before his death last year had suffered the almost complete destruction of his family by the war. M. Boirac seems to have been a very credulous person, and professed his faith in all sorts of occult phenomena, as spiritism, thought transference, vital radiation, and other manifestations of psychic energy. He even contends that this way lies the path of scientific psychology. Such a wreckage of rigorous logical thinking under the emotional stress of the war would be merely pitiful, if it were not for the vicious impulse to credulity which this flood of occult literature from men who will be quoted as authoritative will give to the uncritical masses. From this point of view their influence is most pernicious.

As G. Stanley Hall truly says in that altogether delightful study *A Medium in the Bud,* "Psychic researchers today represent the last potent stand of about all the old superstitions of the past against which science has contended. The next generation will be hardly able to believe that prominent men in this wasted their energies in chasing such a will-o'-the-wisp as the veracity of messages or the reality of a post-mortem existence, which they no more prove than dreams of levitation prove that men can hover in the air at will." If education is to make for straight thinking and right living, specific measures should be taken by teachers even in the elementary school to make their pupils realize how absurd, illogical and superstitious such beliefs are. In the psychological courses to prospective teachers throughout the country a special effort should be made to hold such antiquated medievalism up to ridicule, and thus to work for immunity against this particular disease of the social consciousness.

J. C. B.

NOTES AND NEWS

A group of students of man have formed a "Galton Society" at New York City. The charter members are Madison Grant, George S. Huntington, Charles B. Davenport, C. H. Merriam, William K. Gregory, J. H. McGregor, Edward L. Thorndike, Edwin G. Conklin and H. F. Osborn. It is proposed to elect members from time to time up to twenty-five in number. Dr. Davenport was elected president and Dr. Gregory secretary.

"A Scale for Measuring Ability of Children in Geography in Grades IV–VIII" has been published by H. H. Hahn and E. E. Lackey, of the Wayne State Normal School, Wayne, Nebraska. Answers were obtained to 225 selected geography questions from 1696 pupils in 12 schools. The scale is arranged on the plan of the Ayres Spelling Scale, the questions being divided into twenty-five groups, from the easiest to the hardest, and the percentage made by each school grade on each group being indicated at the top. The authors have promised a more extended account of the derivation of the scale for an early number of the *Journal*.

Mr. Arthur S. Otis has recently issued the printed forms and manual of directions for his Absolute Point Scale, an account of which appears in the present and following issues of the *Journal*. The price will be $10.00 per hundred sets, including the various record sheets. The tests are issued in two forms, A and B, so that the testing can be duplicated at short intervals without any vitiating memory factor.

The Iowa Child Welfare Research Station has issued an individual record blank for recording the achievement of a child in the Terman Scale and the Yerkes Point Scales for six consecutive years. This device, facilitates a rapid survey of the child's intellectual development in successive years.

The Bureau of Educational Measurements and Standards of the Kansas State Normal School has issued an interesting pamphlet setting forth the values of standardized tests to the superintendent, the teacher, and the pupil, and giving a brief description of the more important tests, with prices for the blanks and a reference to the most complete discussion of the test.

The Amsterdam Paidological Association, under the presidency of Dr. G. A. M. Van Wayenburg, has established a new educational journal under the title of "*Kinderstudie*." The journal is issued at irregular intervals, and the subscription price to foreigners is 4.50 florins (approximately $2.00) per volume. The journal will deal with all phases of child development, especially the development of intelligence.

The publishers of "The World Book," Chicago, are issuing an attractive and stimulating monthly 'service bulletin' for the benefit

to teachers. The bulletin contains questions on topics pertaining of school work, and is constructed with the view to arouse an attitude of inquiry toward timely items of general information.

On Friday night, April 26, three of the cottages at Dr. Maximilian P. E. Groszmann's institution for exceptional children, "Watchung Crest," Plainfield, New Jersey, were destroyed by fire. The main building was left uninjured. One of the cottages contained Dr. Groszmann's library and private clinic, and he had the misfortune to lose his entire collection of books, manuscripts, records, notes and correspondence.

The Berkshire Industrial Farm, at Canaan, N. Y., has established a department of clinical psychology under the supervision of Dr. Clinton P. McCord, of Albany. It is planned to have a complete physical, neurological and psychological examination of each boy in the institution, and this will be followed by regular monthly tests thereafter.

Professor Samuel P. Hayes, of Mt. Holyoke College, will appreciate any gifts of psychological literature to help replenish the losses sustained in the burning of the psychological library on December 20.

Among the lecturers from other institutions on the faculty of the summer sesssion of the School of Education, University of Chicago, are William Anton Schmidt, professor of education in the University of Oklahoma, Leonard V. Koos, associate professor of education in the University of Washington, and Edward Herbert Cameron, assistant professor of education in Yale University.

Dr. Thorstein B. Veblen, of the University of Missouri, author of *Theory of the Leisure Class* and *The Instinct of Workmanship*, has been appointed professor of economic institutions in Cornell University.

Miss Katharine B. Graves has been appointed instructor in educational psychology in the Woman's College of Delaware, Neward, Delaware.

Dr. George V. N. Dearborn has been commissioned a firstl ieutenant in the Medical Reserve Corps.

Dr. R. B. Teachout, instructor in psychology in the University of Oregon, has entered the psychological service of the national army.

Lieutenant M. R. Trabue has been promoted to captain and assigned to the officers' training school for psychological examiners at Camp Greenleaf, Fort Oglethorpe, Ga.

VOLUME IX JUNE, 1918 NUMBER 6

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

MONROE'S STANDARDIZED SILENT READING TESTS*

WALTER S. MONROE

Director of Bureau of Educational Measurements, Kansas State Normal School, Emporia, Kansas

The widespread use of the Kansas Silent Reading Tests, devised† by F. J. Kelly is evidence that school men generally feel the need for a simple instrument for measuring the ability of pupils to read. These tests were made available for general distribution in September, 1915, and within a year there was a demand for a duplicate form of the tests or similar tests because in those cities where the tests had been given the pupils were acquainted with the exercises. It was to meet this demand that the derivation of the tests described in this article was undertaken. The author has attempted to introduce certain improvements which will be mentioned below, but he is very conscious of his indebtedness to the originator of the plan upon which the Kansas Silent Reading Tests are based.

The first step in the derivation of reading tests of this type is to secure suitable exercises. The thesis of the writer is that ability to understand sentences and paragraphs depends in part upon one's store of information. Hence, in order to read with understanding in a special field such as arithmetic or geography, it is necessary that one have an acquaintance with this field. In selecting exercises an attempt has been made to restrict them to what we may call the field of general reading and not to include any taken from a particular field, such as arithmetic or geography. Accordingly

*Sample copies of these tests may be obtained by addressing the Bureau of Educational Measurements, Emporia, Kansas. The tests are sold in any quantity desired for fifty cents per one hundred copies.

†An account of the derivation of these tests was published in the Journal of Educational Psychology, February, 1916. See this reference for details of the method used in the derivation of the tests described in this article.

school readers and other books which children read were examined for paragraphs which might be used as a basis for such exercises. To illustrate, in one of the school readers examined this stanza was found,

"O suns and skies and clouds of June,
And flowers of June together,
You can not rival for one hour
October's bright blue weather."

Upon this stanza the following question was used. "Which month does this stanza say is the more pleasant?"

It is realized that in the reading of such a stanza one's purpose may not be to secure information; it may be to enjoy the phrasing and rhythm of the poem, but in so far as one's purpose is intellectual, one should get from the stanza the idea which is called for in the question. There may be other ideas which one should get from this stanza but it seems clear that this is the dominant one.

By examining school readers and other books and by means of the assistance of a class in Educational Measurements, about three hundred exercises were secured. Out of this number 157 were selected to be evaluated. Out of this selected list as will appear later, another selection was made for those which compose the tests.

It is perhaps worthy of note that the greatest difficulty encountered in securing exercises was in finding ones suitable for the lower grades. It was with great difficulty that a sufficient number was secured and the writer believes that they are inferior in quality to those for the upper grades.

The plan* of evaluation used was to have the exercises given to pupils in such a way that the time required and the accuracy of the answers could be noted. For this purpose the exercises were printed on slips of paper with the lines of print approximately the same length as those to which the pupils are accustomed. Ten-point type was used. This perhaps is smaller than the children in the third grade are accustomed to but this criticism does not apply to the grades above the third.

The slips of paper upon which an exercise was printed were stapled together in bundles of five and each labeled with the grade in which the bundle was to be given. The plan of giving the exercises was

*The directions, blanks and several of the details of the plan are not given here. The author will be glad to supply them upon request, or they may be found in the article describing the derivation of the Kansas Silent Reading Tests.

to have the pupils seated so that there were at least five pupils in a row. The pupil in the front seat was given a bundle and began to read it at a signal from the teacher. As soon as he had finished the exercise, he tore off his sheet and passed the remaining papers to the pupil behind him. This process was repeated until the fifth pupil had finished. The total number of seconds and the number of errors were recorded. In determining the grades in which the exercises should be given the assistance of the supervisors of the Normal Training School and of a class of students in Educational Measurements was secured. The exercises were placed in accordance with the concensus of the judgment of these people. It was planned to have each exercise given to three hundred pupils, seventy-five pupils in each of four consecutive grades. Because of the many things which occurred in the fall of 1917 to disturb school work some of those who planned to co-operate in the work found it impossible to give the exercises and consequently this ideal was not attained in many cases.* In a few cases the record reported for certain exercises was practically impossible. Such records were not used in any of the tabulations. However, after excluding these cases the average number of pupils doing each exercise was 186.

For each exercise there were secured and tabulated on a card by grades, the number of pupils to whom the exercise was given, the total number of seconds required and the number of errors made. From these data the average number of seconds required to obtain a correct answer was calculated for each grade. This number re-

*The writer is indebted to the following persons for securing the data upon which the tests are based by giving the exercises in their schools: Supt. D. C. Bliss, Montclair, N. J.; Supt. J. F. Bender, Pittsburg, Kansas.; Supt. W. S. Heusner, Salina, Kans.; Supt. W. A. Walls, Martins Ferry, Ohio; Supt. Earnest C. Witham, Southington, Conn.; Supt. Chas. A. Wagner, Junction City, Kans.; Supt. A. S. Anderson, La Grange, Ill.; Supt. E. B. Gift, Manhattan, Kans.; Supt. A. S. Lyness, Norton, Kans.; Supt. L. W. Fulton, Ellsworth, Wisconsin; Supt. C. A. Yeomans, Minneapolis, Kans.; Supt. J. O. Hall, Hutchinson, Kans.; Supt. H. H. Van Fleet, Aspen, Colo.; Supt. David Gibbs, Meriden, Conn.; Supt. R. E. McCormick, La Crosse, Wisc.; Supt. A. C. Kingsford, Baraboo, Wisc.; Mr. I. J. Bright, Topeka, Kans.; Supt. M. E. Pearson, Kansas City, Kansas; Supt. S. H. Holmes, New Britian, Conn.; Prin. A. S. Philips, Denver, Colo.; Asst. Supt. J. W. Studebaker, Des Moines, Ia.; Supt. F. H. Prout, Chillicothe, Ohio; Supt. F. W. Travell, Ridgewood, N. J.; Supt. P. S. Berg, Dickinson, N. Dak.; Mrs. Mary A. Davis, Claremore, Okla.; Mr. George Melcher, Kansas City, Mo.; Supt. Frederick B. Knight, Danvers, Mass.; Supt. E. C. Fisher, Rock Island, Ill.; and Mr. W. S. Dakin, Hartford, Conn.

presents the "difficulty" of the exercise. The "weighted difficulty" was obtained by the method described in the following paragraphs.*

Since the exercises were given in the successive grades it was possible to compare the reading ability of fourth grade pupils with that of third grade pupils when working on the same exercises. In Table I there is given for the several grades by pairs the number of pupils who took the same exercises, the total number of seconds required for these pupils to do the exercises, the number of errors made by the pupils and the average number of seconds per correct answer. This last item is obtained by subtracting the number of errors from the number of pupils and using this remainder for the divisor into the total number of seconds. Third grade pupils required on the average 143 seconds for each correct answer. When doing the same exercises it required only an average of 74 seconds for fourth grade pupils to secure a correct answer. The quotient of the reciprocals of these two numbers (74:143) represents the ratio of the silent reading ability of third grade pupils to silent reading ability of fourth grade pupils.

From the facts of Table I it is possible to calculate ratio of the silent reading ability for any grade to the silent reading ability of the third grade. These ratios are given in Table II. They are the factors by which the average number of seconds required for pupils of any grade to do an exercise correctly (or the "difficulty" of an exercise) must be multiplied to reduce the difficulty to the basis of third-grade silent reading ability. This product (the average number of seconds required to secure a correct answer times the ratio for that grade) is called the "weighted difficulty."

In calculating the average difficulty of an exercise it was thought wise not to use the "weighted difficulty" of a grade in which the exercise was not to be used in the test. For example, if an exercise had been given originally in grades 3, 4, 5, and 6 and it was to be used only in grades 3, 4, and 5 the "weighted difficulty" for the sixth grade was not used in determining the average difficulty.

From the average difficulty the comprehension score has been calculated by simply dividing the average difficulty by 81. The

*Tables were prepared showing the number of pupils in each grade reading each exercise, the total number of pupils reading each exercise, the weighted difficulty for each grade, the average difficulty, comprehension score, and the place to which the exercise was assigned. Owing to limitations of space these tables cannot be reproduced here.

choice of 81 as a divisor has no significance. It was chosen because it happened that this was the lowest average difficulty of a number of exercises which were used in a preliminary tabulation.

In choosing exercises for the tests an effort was made to reject (1) all exercises considered unsatisfactory by those who co-operated in giving exercises and (2) those exercises for which the weighted difficulties in the successive grades were not approximately equal. In every case an exercise was excluded if commented upon adversely by those co-operating. It was not possible to reject in every case those exercises whose statistical behavior was not wholly satisfactory but it is believed that no exercise was included which is open to serious criticism on this basis.

In a report on the use of Kansas Silent Reading Tests* the writer has shown that a score on these tests measures the comprehension rather than the rate of reading. The method used for determining the comprehension value described above is essentially the same as that used by Kelly in the case of the Kansas Silent Reading Tests. For this reason it has seemed wise to give each exercise a rate score. The rate score has been determined by counting the number of words in an exercise and dividing the total by five. Five was used as a divisor because five minutes are allowed for giving the tests and thus the pupils' rate score will be the number of words read per minute. It should be noted, however, that since the material read is not uniform in difficulty this rate score can have only a very general meaning.

The number of exercises was sufficient to make it possible to construct three forms of Test I and Test II and two forms of Test III. In Table III the comprehension values of the exercises which compose the different forms of these tests are given. It was not possible to choose exercises for duplicate forms with exactly equal values. Just how much the slight differences will affect the scores can be determined only by actual trial. At the time when this report was written this had not been done. It should be remembered that considerations other than the comprehension values entered into the choice of exercises for the several forms of the tests. It was necessary to give consideration to the length of the exercises, because the tests were printed on a four page folder of fixed size. This is the reason why the number of exercises for the different

*Walter S. Monroe. *Report on the Kansas Silent Reading Tests With Over One Hundred Thousand Children.* Journal of Educational Psychology, December, 1917.

forms of Test I is not the same. Attention was also given to the general character of the exercises. Test II, Form I is reproduced to illustrate the nature of the series of tests.

In Table IV median scores are given which may be used as tentative standards. The median scores for grades 9 to 12 are based on returns from only a few cities and in addition the number of pupils in each grade is very small. Thus the medians for these grades are probably not reliable standards. However, the medians for grades 3 to 8 are based upon a sufficient number of cases so that it is very probable that additional data would make few changes. The tests on which these medians (3 to 8) are based were given between April first and the close of the school year. Hence they may be considered as end of the year standards.

It should be noted that the decrease in the rate of reading from the fifth to the sixth grade and from the eighth to the ninth is to be expected because at these points different tests are used. The exercises of the tests are more difficult and require more time for reading.

The comprehension medians approximate the learning curve except from the fifth to the sixth grade and from the eighth to the ninth grade. The drop from the eighth to the ninth might be explained by the relatively small number of cases in the high school grades, but the same explanation will not apply to the other case. It appears, therefore, that although an effort was made to have the exercises of the several tests evaluated in terms of a common unit, the units in Test II are larger than those in Test I and those in Test III are larger than those in Test II. The ratio of these units must be determined in order that progress may be judged at the points where the tests are changed.

In comparison with the Kansas Silent Reading Tests it is believed that these tests possess the following points of superiority.

1. The exercises are more uniform in character. They are confined to what we may call the field of general reading.

2. The exercises have been more reliably evaluated. Kelly had the exercises given to only 60 pupils. The exercises of these tests were given on the average to 186 pupils. In the second place, the exercises were printed and not mimeographed. It should also be noted that the pupils to whom the exercises were given were not confined to any particular state or section of the country.

3. The comprehension values of the exercises are more uniform. They range only from 1 to 10.1 with one exception and the value of that exercise is 18.5.

4. Duplicate forms of the tests consisting of exercises with approximately equivalent values have been provided.

STANDARDIZED SILENT READING TEST, TEST II, FORM 1

For Grades 6, 7 and 8

No. 1

Rate Value 5 Comprehension Value 2.1

Oil floats on water because it is lighter than water. Milk mixes with water because milk and water are of equal weight.

Which is lighter, oil or milk?....................

No. 2

Rate Value 7 Comprehension Value 2.1

At evening when I go to bed
I see the stars shine overhead;
They are the little daisies white
That dot the meadow of the night.

What are the little white daisies of the night?

....

No. 3

Rate Value 12 Comprehension Value 2.7

The dog lay down. The rooster flew to the top of a tree and the cat climbed to one of the branches. Before they went to sleep the rooster saw a light in the forest. He called to his friends.

Where was the light the rooster saw? Draw a line around the word that tells where.

sky house barn wagon forest

No. 4

Rate Value 7 Comprehension Value 2.8

O sun and skies and clouds of June,
And flowers of June together,
You can not rival for one hour
October's bright blue weather.

Which month does this stanza say is the more pleasant? ....

No. 5

Rate Value 8 Comprehension Value 3.1

They rested and talked. Their talk was all about their flocks, a dull theme to the world, yet a theme which was all the world to them.

What do you suppose was the occupation of these men?

carpenter doctor merchant shepherd blacksmith

No. 6

Rate Value 4 Comprehension Value 3.3

Africa is smaller than Asia, and North America is not as large as Africa. Which is smaller, Asia or North America? ....

No. 7

Rate Value 11 Comprehension Value 3.2

The fly does not care whether it is king or clown whom he teases. He has no work to do—no instinct to obey. The earthworm has his digging; the bee her gathering and building; the spider her cunning network; the ant her treasury. They are slaves.

Is the fly a free creature or a slave? ....

No. 8

Rate Value 15 Comprehension Value 3.7

He said to his friend: "If the British march
By land or sea from the town tonight,
Hang a lantern aloft in the belfry-arch
Of the North Church tower, as a signal light,
One if by land and two if by sea;
And I on the opposite shore will be
Ready to ride and spread the alarm."

If the British came by land, how many lights did the man see in the church tower?

....

No. 9

Rate Value 16 Comprehension Value 3.8

It was cold, bleak, biting weather; foggy withal; and he could hear the people in the court outside go wheezing up and down, beating their hands upon their breasts and stamping their feet upon the pavement-stones to warm them.

The author has attempted to give you a picture in this paragraph. After reading the paragraph, if you think it is a picture of comfort and pleasantness, underline the word hear; if of cheerlessness and dreariness, underline bleak.

hear wind bleak cold

No. 10

Rate Value 13 Comprehension Value 4.1

It is a beautiful moonlight night. Mary and Kate are invited to a party. Kate promised to come by for Mary. Mary has hurried to get ready and is impatiently waiting for Kate. When it was too late to go, Kate's mother telephoned that Kate could not go.

Underline the word below that most nearly describes how Mary now felt.

happy comfortable sad excited

No. 11

Rate Value 10 Comprehension Value 4.4

The boy stood on the burning deck,
Whence all but he had fled;
The flame that lit the battle's wreck,
Shone round him o'er the dead:
Yet beautiful and bright he stood,
As born to rule the storm.

Underline the word that best describes the boy.

cowardly mischievous brave young good

No. 12

Rate Value 11 Comprehension Value 4.9

Aladdin's uncle said: "I will take a shop and furnish it for you." Aladdin was delighted with the idea, for he thought there was very little work in keeping a shop. He liked that better than anything else.

Underline the word below that tells us what kind of a boy Aladdin was.

industrious ambitious active lazy honest

No. 13

Rate Value 14 Comprehension Value 5.2

Not far from Greensburg is a little valley, among the high hills. A small brook glides through it, with just murmur enough to lull one to repose; and the occasional whistle of a quail, or tapping of a woodpecker, is almost the only sound that ever breaks in upon the uniform tranquillity.

What kind of a picture do you get from reading the above paragraph?

disorder activity noise calmness confusion

No. 14

Rate Value 13 Comprehension Value 5.4

The soldier crawled out of the trench, where he had spent the night. He was covered with mud from head to foot, and almost frozen. He looked around at his companions. What a miserable lot they were! This, then, was the glorious war told about in the papers.

Underline the word below that tells how the soldier felt.

happy patriotic brave angry downhearted

TABLE I

Showing Relative Ability of Pupils in Successive Grades Doing Same Exercise

Grade	III	IV	IV	V	V	VI	VI	VII
Number of Pupils	1,801	1,793	2,449	2,421	3,526	3,533	3,774	3,839
Total Seconds	146,962	95,078	138,514	87,751	157,620	136,143	155,237	160,859
Errors	772	511	845	588	1,193	1,145	1,327	1,028
Ave. No. Sec. per Correct Answer	143	74	86	48	68	57	63.5	57

TABLE I—*Concluded*

VII	VIII	VIII	IX	IX	X	X	XI	XI	XII
2,407	2,378	1,283	1,302	1,302	1,317	1,253	1,253	1,008	1,070
101,761	89,973	55,981	51,723	61,086	62,657	58,511	53,875	45,545	42,460
618	552	393	372	520	476	451	387	339	317
57	49	63	56	78	75	73	63	68	56

TABLE II

Ratio of Silent Reading Ability of Pupils in the Several Grades

Grade....	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Ratio....	1.00	1.92	3.44	4.16	4.65	5.40	6.06	6.29	7.10	8.85

TABLE III

Showing Comprehension Values of the Exercises of the Different Forms

Ex. No.	Test I Form I	Test I Form II	Test I Form III	Test II Form I	Test II Form II	Test II Form III	Test III Form I	Test III Form II
1.	1.3	1.1	1.0	2.0	2.0	2.2	2.8	3.5
2.	1.3	1.1	1.2	2.1	2.8	2.3	3.5	3.6
3.	1.3	1.3	1.3	2.7	2.7	2.6	4.8	4.0
4.	1.4	1.4	1.4	2.8	2.8	2.7	4.9	4.4
5.	1.4	1.4	1.5	2.8	3.2	3.3	5.0	4.8
6.	1.5	1.5	1.5	3.0	3.3	3.2	5.6	5.4
7.	1.7	1.7	1.7	3.0	3.7	3.7	5.6	5.7
8.	1.8	1.8	1.8	3.4	3.7	3.8	5.9	5.8
9.	2.0	2.3	2.0	3.4	3.8	3.9	6.6	7.6
10.	2.3	2.1	2.3	3.6	4.0	4.0	6.7	8.9
11.	2.5	2.4	2.3	4.1	4.1	4.6	7.7	9.2
12.	2.7	2.5	2.4	4.3	4.6	4.7	18.5	10.1
13.	2.9	2.6	2.9	4.4	4.7	5.2		
14.	2.9	2.8	3.0					
15.	3.0	3.0						
16.	3.5							

TABLE IV

Standards—Median Scores

Grade	III	IV	V	VI	VII	VIII	IX	X	XI	XII
No. of Pupils	5402	5519	5221	4734	3752	2998	349	355	282	162
Rate	57	83	91	96	98	108	87	81	88	89
Comprehension	8.7	15	20	21	25	28	22	25	26.4	27

PSYCHOLOGICAL ASPECTS OF LANGUAGE

GEORGE C. BRANDENBURG
Purdue University, Lafayette, Ind.

Purpose and Method

Purpose. The investigation discussed in this article was undertaken for the purpose of contributing some data upon the question of the relation of language to general intelligence. It is hoped that the material presented may aid in a small degree in answering the following questions: What is the relation between the extent of a pupil's vocabulary and his scholastic attainments? What correlation exists between accuracy and precision in the knowledge and use of words, and vocabulary range? What are the chief factors in the acquisition and use of words? To what extent does present-day school room practice facilitate the pupil's linguistic development?

Method. From a series of preliminary investigations it was decided that the following principles were valid for a vocabulary test: (1) The best test as to whether a word is sufficiently well known by a pupil to be credited to him is his ability to use it correctly in a sentence. (2) The range of vocabulary varies so much within any group that a pupil must be tested upon approximately the entire list of words used, and this list must range from the most simple to the most difficult. (3) The list of test words should be short enough to be completed by the average pupil at one sitting of thirty or forty minutes. Accordingly the final test was arranged containing 200 words, one for every 140 in Webster's Academic Dictionary. The words were selected at equal intervals throughout the dictionary. The test was printed in the form of a four page folder, the words being arranged as nearly as possible in order of difficulty, and each pupil was furnished with a copy. The test was given to 2000 pupils of 68 different classes from the second to the twelfth grades, in 16 schools of six different school systems located in Wisconsin, Iowa, Missouri and Colorado. The places in which the tests were given vary in size and character from an agricultural town of 1200 people to a small modern city of 30,000. The author supervised personally the giving of 800 of the tests; the remainder were given under the direction of the various principals of these schools. The tests were all given during the month of May so that the results show the rankings and averages of the various

grades at the end of the school year. The instructions to the pupils which were placed at the top of the first page of the folder were as follows: "In the blank space after each word that you know, write a sentence using the word correctly. Place a cross before each word that you do not know. For seventh grade pupils and above omit the first 25 words." On the last page of the folder were spaces for the pupils' grades in English, manual training, oral expression and average scholarship, with a request that these grades be supplied by the teacher.

With respect to scoring results any such test is, to a certain extent, defective, since it is not always possible to decide from a pupil's use of a word in a sentence whether he knows the word or not. The principle which was followed in grading the papers was that the attempted use of a word is an indication in itself that the word is known, and unless the use in the sentence was such as to indicate that it was not known, the pupil was given credit for it.

Grade Averages and Variations in Vocabulary

In a study somewhat similar to the present one Kirkpatrick (7) attempted to find the average size of the vocabulary for each grade, and, while the method used was different, the results did not differ materially from those found here. His averages were as follows: Grade II, 4480; Grade III, 6620; Grade IV, 7020; Grade V, 7860; Grade VI, 8700; Grade VII, 10660; Grade VIII, 12000. Comparing these results with those in Table I, it will be seen that the increase in size of vocabulary from year to year appears much more uniform here than Kirkpatrick found it to be. The averages for Grades II and III and for those above the eighth in Table I cannot be regarded as reliable on account of the small number of pupils tested. It will be noted that the average of the three tenth grades tested is larger than either the eleventh or twelfth grade average. No explanation of this can be offered except that in the case of two of the tenth grades the principals remarked that they were unusually strong classes. The average gain from year to year for Grades II to VIII is approximately 1400 words. It appears that the gain in the seventh grade is greater than that for any other year, being 371 above the average. This is seen to be the case also in Kirkpatrick's results, disregarding Grades II and III which are admittedly unreliable. It was found by Dr. Jones, in a study of pupils' compositions that the greatest gain in the number of words

used by pupils of one grade over those of the preceding grade was in the seventh grade. It is quite likely that this unusual increase in linguistic ability in this grade is due to the fact that a large number of dull pupils drop out of school at the end of the sixth year. This

TABLE I

Average Size of Vocabulary for the Various Grades as Determined by All the Tests Which Were Completed

Grade	No. of classes	No. of pupils	Av. vocabulary	Gain over preceding year	A. D. of classes	Per cent deviation
II	1	22	4000			
III	3	78	5429	1429	453	6.5
IV	8	228	6887	1458	453	6.5
V	10	245	8207	1320	717	8.7
VI	13	378	9613	1406	754	7.8
VII	11	300	11445	1832	581	5.0
VIII	10	255	12819	1374	779	6.0
IX	3	72	13504	685		
X	3	71	15340	1736		
XI	3	71	13974			
XII	3	41	14975			

A. D., average deviation.

would seem to explain the fact that there is a smaller percentage of variation among seventh grades than among the classes of any other grade. Some rather surprising facts are discovered when we compare the grades of different schools: The highest fourth grade had an average vocabulary of 8806, being larger than the average of the fifth grades and almost as large as the average of the sixth. The best eighth grade with a vocabulary of 14218 apparently had a better knowledge and command of words than the poorest senior high school class, and almost as good as the average of the three senior classes. These cases are typical of the overlapping in vocabulary ability throughout. In a later section dealing with the causes of individual variation in vocabulary an attempt will be made to explain these variations in grade averages, since whatever factors operate to cause individual differences would also tend to cause differences between grades.

Averages for Boys and Girls

The differences between boys and girls in the matter of vocabulary do not appear to be great as will be seen from the following table, and not uniformly in favor of either.

TABLE II

Grade	Boys	Average	Girls	Average
III	38	5527	40	5372
IV	126	6546	102	7308
V	114	8165	131	8320
VI	210	9886	168	9327
VII	146	11363	154	11472
VIII	113	12914	142	12627
IX	38	13357	34	13659
X	45	15112	26	15738
XI	41	14294	30	13352
XII	21	15634	20	14273
Average	892	10232	847	10156

The girls average 76 words less than the boys, an entirely insignificant difference. The boys rank higher in five grades, while the girls are higher in the other five.

The Relation of School Abilities to Range of Vocabulary

The grades of approximately 1500 pupils were returned on the test sheets sent to teachers, about 500 papers being returned without the grades. For about 200 pupils the grades in average scholarship and language were given in figures, the remainder being in letters or words. It was not expected that the grades in oral expression would be very reliable since teachers, generally, pay little attention to grading this ability. It was thought, however, that the estimates thus secured would aid in discussing the proposition that there are many persons who have large vocabularies and are still unable to give clear expression to their thoughts.

Language Grades and Vocabulary Ability. Table III contains a summary of the results of the test in the classes in which the language grades were returned in figures. The vocabulary index is the number of words correctly used from the 200 word list. The coefficient of correlation is represented by r, and was calculated by the Pearson formula. The average deviation is indicated by A. D.

TABLE III

			Vocabulary Indices				Language Grades				
School	Grade	No. pupils	Av.	Highest	Lowest	A. D.	Av.	Highest	Lowest	A. D.	r.
X	IV	26	48	65	26	6.1	83.8	91	72	4.0	.73
II	V	17	60.7	77	46	6.6	80.3	91	70	5.2	.90
IX	VI	35	72	89	43	9.7	83.7	95	70	6.6	.54
X	VI	20	59.2	70	43.6	6.8	79.5	88	67	5.5	.85
VI	VII	40	80.5	109	48	7.8	80.5	95	60	4.9	.82
XI	VII	18	84	106	66	7.7	85	95	70	6.6	.66
I	VIII	26	84	114	63	12.8	83.4	91.5	72	3.6	.72
XV	X	21	110.5	128	82	11.7	86.5	95	70	5.6	.81

In this table two features stand out prominently: (1) The marked deviations from the average in vocabulary ability and the absence of such deviations in the language grades; (2) the high degree of correlation between the two abilities considered. The latter is of considerable significance to the teacher since it indicates that she could by a vocabulary test estimate the language ability of her pupils in an objective and impersonal manner, accurately and easily. The fact that all the pupils of a certain group, as for example, Grade VIII, School I, are using the same text book, doing the same outside reading, are expected to write the same sort of compositions, in short, are subject to the same linguistic requirements, when certain individuals in the group understand and are capable of using almost twice as many different words as are certain others, points strongly toward the conclusion that there is something fundamentally inadequate in the basis for the grouping.

It is somewhat surprising, considering the various factors tending to produce abnormal results, which are usually present in giving such a test, that the degree of correlation should be so high. The lowest coefficient is .54, the highest .90, while the average of the eight coefficients is .76.

Table IV shows the relation between the vocabulary abilities and the language grades of 1000 pupils whose grades were given in letters or words.

TABLE IV

Language grade	No. pupils	*Distribution on Basis of Vocabulary Ability*							
		Excellent		Good		Fair		Poor	
		No.	%	No.	%	No.	%	No	%
Excellent	146	86	59	56	38.3	4	2.7	0	0
Good	459	55	12	295	64.3	94	20.5	15	3.2
Fair	291	5	1.7	100	34.4	144	49.4	42	14.4
Poor	104	0	0	6	5.8	43	41.3	55	52.8

In order to make clear the method by which these correlations were worked out let us consider a group in which the teacher had marked 5 pupils E; 22, G; 8, F and 1, P. Assuming this apportionment for the group, we then divided the class as to vocabulary scores in the same proportions. We then find that three of the E pupils in language fall within the E division in vocabulary, and two in the G group; that 16 of the G pupils in language fall within the G group in vocabulary, two in the E group, and four in the F group; that three of the F pupils in language fall within the F group in

vocabulary, four in the G, and one in the P group; that the one P pupil falls in the F group in vocabulary. The scores of the 1000 pupils were distributed relatively to the grades in this manner.

The fact that the correlation is higher in the "Excellent" and "Good" divisions probably indicates that teachers grade more accurately here than in the "Fair" and "Poor" groups. That is, they probably tend to raise the grades of fair or poor pupils above what they should be.

Oral Expression and Range of Vocabulary. When the grades in oral expression are compared with vocabulary abilities we find about the same relationship as existed between language grades and vocabulary ability. The average tendency of grades in oral expression

TABLE V

Oral Expression grades	No. pupils	*Distribution on Basis of Vocabulary Ability* Excellent		Good		Fair		Poor	
		No.	%	No.	%	No.	%	No.	%
Excellent	180	110	61.1	58	32.2	12	6.7	0	0
Good	418	62	14.8	283	67.7	62	14.8	11	2.6
Fair	260	7	2.7	75	28.9	128	49.2	50	19.2
Poor	142	0	0	7	4.9	52	36.7	83	58.4

to fall within their proper or corresponding groups in vocabulary ability is seen to be slightly above 59%. It will be noted that no pupil who was excellent in vocabulary was poor in oral expression and that no pupil who was poor in vocabulary was graded excellent in oral expression. Here again we have evidence of the better grading by teachers in the "Excellent" and "Good" groups. From the material presented here we are forced to the conclusion that there are relatively few persons who have large vocabularies and are not good in oral expression. The fact that about 5% of the pupils who were marked poor in oral expression are in the good division in vocabulary indicates that there probably are such individuals.

Range of Vocabulary and Average Scholarship. The average scholarship grades, in most cases, represent the teachers' estimates of the quality and quantity of pupils' achievements in all their studies except manual training. These grades were given in figures for only five classes. It will be noted that the coefficients of correlation vary from .39 to .85, the average of the five being .63.

TABLE VI

School	Grade	No. pupils	Vocabulary Indices Av.	Highest	Lowest	A. D.	Average Scholarship Grades Av.	Highest	Lowest	A. D.	r.
X	IV	26	48.2	65	26	6.1	81.5	91	75	4.1	.85
IX	VI	35	72	89	43	9.7	80	90	65	6.9	.60
XI	VII	18	84	106	66	6.6	84.7	95	70	5.9	.64
VII	VII	20	84.4	118	65	16.3	79	84	74	1.9	.39
XV	X	21	110.5	128	82	11.7	87	94	80	4.3	.70

TABLE VII

Average Scholarship Grade	No. of pupils	Distribution on Basis of Vocabulary Ability: Excellent No.	%	Good No.	%	Fair No.	%	Poor No.	%
Excellent	163	93	57	65	39.8	5	3.1	0	0
Good	427	60	14	270	63.2	82	19.2	15	3.5
Fair	277	10	3.6	82	29.6	140	50.6	45	16.1
Poor	133	0	0	12	9.	48	3.6	73	55.

It would seem from Table VII that there is as high a correlation between average scholarship and size of vocabulary as there is between oral expeession and vocabulary ability. The average per cent. of exact correspondence in groups is approximately the same in both cases, a little above 56 per cent.

In view of the above results we may safely say that the opinion sometimes ventured, that there are certain individuals who have a large fund of words at their command and are very proficient in the use of words generally but are mentally weak and inefficient, is by no means well founded. The only other alternative is that average scholarship grades are not indcative of mental capacity; it would be difficult, indeed, to defend such a proposition.

Manual Training Grades and Range of Vocabulary. The theory that motor ability and verbal capacity do not tend strongly to coincide is supported by the following table.

TABLE VIII

Manual Training Grade	No. pupils	*Distribution on Basis of Vocabulary Ability*: Excellent No.	%	Good No.	%	Fair No.	%	Poor No.	%
Excellent	66	23	34.8	35	53	8	12.1	0	0
Good	248	39	15.7	179	72.2	20	8.	10	4.
Fair	70	4	5.7	31	44.3	32	45.7	3	4.2
Poor	36	0	0	6	16.7	15	41.6	15	41.6

It will be noted that more of the E pupils in manual training come in the G rank in vocabulary than come in the E rank; that about as large a percentage of the F pupils in manual training come in the G rank in vocabulary as come in the F rank; while just as many of the P pupils in manual training come in the F

group in vocabulary as come in the P group. It is significant, however, that no E pupil in manual training is found in the P group in vocabulary.

As a check upon these results a personal study was made of a group of 48 high school pupils. After securing the instructor's grades upon these pupils in manual training, they were given the vocabulary test, and the correlation between the two abilities was found to be .21. It was found by inspection of the details that the pupil who ranked highest in vocabulary was marked 2% below the average in manual training, and that the pupil who ranked lowest in vocabulary was graded 5% above the average of his class in manual training. On the other hand, all the pupils who had marks as low as 60% in manual training were found in the lower half of the group in vocabulary ability.

Difficulties and Errors in the Use of Words

The list of test words is here given arranged in order of their relative difficulty upon the basis of results of the test.

Orange	heart	seek	direct	charity
cup	hour	balance	favor	reclaim
penny	worse	beetle	shaft	unfortunately
light	stump	kernel	49 psalm	crinkly
frog	grain	skeleton	text	insure
storm	report	*	revolution	landman
week	roast	failing	gill	lonesomeness
twenty	important	somerset	partition	ferry
flat	shiny	39 childhood	ye	69 bower
third	remain	rate	hippopotamus	lubber
place	foamy	labor	betrayer	bred
square	absent	diving	triangle	inclined
be	barehead	sup	morsel	everlastingly
east	29 camera	content	59 foremost	vail
frame	snuff	gaily	crackle	numbness
82 downcast	polygamist	potent	circularity	anastrophe
observation	picklock	interposer	de capo	epicarp
inkhorn	grievance	municipality	saline	papilionaceous
sycamore	abate	complexity	quayage	mellifluous
pavilion	spawn	proportionateness	unbias	palestra
pleurisy	winebibber	tiffany	marline	arrogate
discrown	slightness	putrid	tocsin	apotheosize
bloat	phenomenal	debauch	preclusive	embracery
determination	messieurs	rotunda	transcendentalist	milt
spinster	aquarium	dissension	immaterialism	templet
copperhead	centiped	coequal	gastronomist	tump
external	restriction	tactical	anodyne	decortication
savory	banker	vestige	trivet	homunculus
buckboard	Finnish	caprice	conservancy	vitrifacture
surpassing	misgotten	proboscis	comatose	pretermit
sickishness	err	cosmetic	carpologist	deltoid
92 upholster	jade	cajolery	monetization	algoid
reexport	leniency	hyperbolical	tatterdemalion	stannary
alternate	layering	whitleather	iridium	endemical
dromedary	abduction	attainability	wady	nereis
jugular	deplorably	binnacle	malleate	ryot
nobby	venom	sere	henna	magian
infidel	overmaster	scholasticism	materia medica	gymnotus
Olympic	humanitarian	effusiveness	bogle	acescent
agate	subsist	clearstarch	autonomic	adytum

*Through an error in printing the word *skeleton* occurs twice.

The number of words which any grade should know from the list is found by dividing the average vocabulary of the grade by 140. Accordingly grade II should be able to use 29 words; grade III, 39; grade IV, 49; grade V, 59; grade VI, 69; grade VII, 82; grade VIII, 92.

None of the last 25 words were used correctly by any pupil; that is, 12½ per cent. of the words in Webster's Academic Dictionary are entirely foreign to high school pupils. There were, naturally, many errors made in the attempted use of words which were not due to real inability to use the words. The following examples taken from the test sheets are suggested as explanatory of such mistakes: (1) Those due to imperfect visual imagery; *e. g.*, "They sold the lubber from the whale," "Psalm trees grow here." (2) Errors due to imperfect auditory imagery; *e. g.*, "Our you a boy?" "I will potent I am dead." (3) Errors due partly to both imperfect auditory and visual imagery; *e. g.*, "The horse is a cosmetic animal," "I pulled the spinster out of my foot." (4) Errors due partly to confused auditory and visual imagery, and partly to lack of familiarity with the word; *e. g.*, "She made her debauch at the party" "It is an autonomic gun." Of course, there is no hard and fast line between these four types of errors. It may seem strange that any errors should be due to auditory imagery since the words were not pronounced to the pupil; but in such cases as the confusion of "our" and "are," since there is little similarity in the appearance of the two words, it seems probable that the pupil images the sound of the word and thus confuses it with a word similar in sound.

Possible Factors in the Acquisition of a Vocabulary

Of the two factors operating in a greater or less degree in practically all mental processes, native endowment and environment, it is generally held that the latter is more important so far as the acquisition of a good vocabulary is concerned.

The most prominent environmental factors are, the home, the community, reading facilities, and travel. An effort was made by means of a personal study of two groups of children to determine what part each of these factors had played.

Home. It is apparent at the outset that the task of grading homes as one grades examination papers is an exceedingly complicated one. Each home represented in the first group of children, called Group A, was visited in order to arrive at an estimation of the home

life on the basis of such characteristics as, occupation of parents, economic status, books and magazines, and general intellectual tone. The homes were then graded as "Excellent," "Good," "Fair," and "Poor." Of the 32 homes, 7 were graded E. Two of the seven children from these homes had vocabularies at, or below the average for the group, while four were in the highest 25 per cent. of the group. In 82 per cent. of the cases the correspondence between groups was close; that is, the pupils from "Excellent" homes had "Excellent" or "Good" vocabularies and those from "Poor" homes had "Poor" or only "Fair" vocabularies. In 18% of the cases there was entire lack of correspondence. For example, the pupil having the third highest vocabulary in the group came from a "Fair" home and the pupil who ranked eighth in vocabulary came from the poorest home in the whole group.

With reference to Group B the fact that all but two of the 25 children were from homes of professors or instructors in the University is sufficient characterization of the homes from which they came. Combining this with the fact that eight of the pupils in the group had vocabularies below the average of their respective grades, we see that the correspondence here between the character of the home and size of vocabulary cannot be close.

So far as the evidence here goes, it points to the conclusion that, while the home is generally an important factor in the acquisition of a good vocabulary, it is not in all cases the determining factor.

Community. With the community, as with the home influence, we are dealing with a number of factors instead of with a single one. Probably the school is the most important element, for the child, of community influence. Since, however, it operates upon all the pupils within a grade in a uniform manner, it would tend to eradicate rather than produce individual differences. The size of the community in which pupils lived was investigated in a number of cases. It was found that the pupil who ranked highest of all the pupils tested, had always lived in a city of several thousand population; while the eighth grade pupil ranking highest of the eighth grade pupils had always lived in a small town. This was typical of the conditions and relations. In other words, the size of the community does not appear to be an important matter so far as vocabulary is concerned.

It might be supposed that the general intellectual tone of a community would have much to do with a child's vocabulary, but in-

asmuch as there were in Group B, in a highly intellectual community, a number of children with inferior vocabularies, it would seem that this factor is not, unless in a very general way, significant.

Reading. It can hardly be doubted that both the quality and quantity of a child's reading has much to do with his verbal ability. Frequently, in explaining a pupil's remarkable performance in the vocabulary test, the teacher would remark that he or she was a great reader. Through a study of group B in which the time spent in reading, the kind of reading, the reading and the enjoyment in reading, were considered, the following facts were ascertained:

(1) The boys read a greater variety than did the girls.

(2) The girls read more than did the boys.

(3) The boys read more humorous material, the girls more light fiction.

(4) No pupil ranked high in vocabulary who was not a great reader.

(5) No pupil ranked high who did not read a variety of literature.

(6) Some who were great readers ranked low in vocabulary.

(7) No one ranked low in vocabulary who read a great variety of literature.

It thus appears that there is a close relation between a child's reading and his command of words; but it is not so much the quantity as the quality that counts.

Travel. The published discussions of childrens' vocabularies stress the matter of travel as a productive source of new words. Travel is so invariably associated with a good home that it was impossible to determine even its approximate effect. It is reasonable to suppose that during the early stages of linguistic development, travel is an important agency in enlarging the vocabulary. After the child begins to read, however, if he has access to an abundance of good literature, it is doubtful whether he gains many words from travel that he would not get from reading. Yet it is not unlikely that his word images become more rounded out, more perfectly defined, and more definitely fixed through travel.

Accuracy in Recognition of Words and Size of Vocabulary. The recognition of a word, no doubt, depends largely upon the distinctness with which the image of the word has been created in the mind; hence an individual in whose mind clear and accurate images are constantly formed will, other things being equal, build up a vocabulary faster than one whose mentality does not function so precisely

in this respect. In order to test the validity of this theory the percentage of error in 500 test papers, selected by chance, except that boys and girls were equally represented, was calculated. Table IX gives the distribution of the 250 boys and 250 girls according to the per cent. of error in the use of words.

TABLE IX

	Less than 10% of error	From 10% to 24% of error	From 25% to 49% of error	Over 49% of error
Boys...........	7.6%	51.6%	40.4%	.4%
Girls...........	18.4%	51.2%	28.8%	1.6%
Total...........	13.0%	51.4%	34.6%	1.0%

The results here lead one to believe that girls are more accurate in the recognition and use of words than are boys. Now if the tendency to greater accuracy on the part of girls is a sex trait we should expect the girls to have larger vocabularies than the boys. We have seen that they do not. The explanation of this is probably found in the differences in environmental conditions of boys and girls. The boys having a wider range of contact with things, persons, and situations, experience a greater number and variety of images together with the various verbal symbols of these images. At the same time, on account of a greater number and variety, they necessarily experience each of these images with its verbal symbol a less number of times than they would if they came into contact with fewer things and with each thing a greater number of times. Hence the images are less accurately fixed and defined. This would be true of experience gained in reading as well. We have seen that boys read less than do girls, but that they read a greater variety of literature; *i. e.*, they come into contact with more different words in reading than do the girls, but necessarily with each fewer times. Consequently they tend to be more uncertain of their words than do the girls. Environmental conditions, then, rather than innate qualities, seem adequate to explain the greater accuracy of the girls.

Coefficients of rank correlation between size of vocabulary and accuracy in recognition of words were calculated for five groups of pupils and the results varied from .18 to .49. The fact that girls are more accurate, as is shown in Table IX, and that boys have slightly larger vocabularies, as is indicated in Table II, accounts for the low degree of correlation. The very fact, however, that

there is a fair degree of correlation between the two capacities, in spite of the influence of environment, tends to substantiate the theory that the vocabulary of an individual does depend in a considerable degree upon the readiness and accuracy with which he forms verbal images.

General Intelligence and Verbal Ability

The three views extant, with regard to the relation of thought to language are summarized by Professor Dewey (5) thus: "First, that they are identical; second, that words are the garb or clothing of thought, necessary not for thought but for conveying it; and third, that while language is not thought it is necessary for thinking as well as for its communication." The first view has had little support among eminent thinkers since the time of Müller. One of the foremost advocates of the second is Professor Preyer, who, after making an extensive study of the mental development of his child, Axel, concluded that the child did not need "words or looks or gestures or any symbol whatever in order to arrange in time and space the sense of perceptions." The third view is the one to which Professor Dewey subscribes and the one which is most generally held.

One of the first difficulties which confronts one in a discussion of the question is the lack of a generally accepted definition and understanding of the term, intelligence. Stern's definition, "Intelligence is a general capacity of an individual consciously to adjust his thinking to new requirements; it is general mental adaptability to new problems and conditions of life," (17, p. 3) is probably the most satisfactory statement for the present state of knowledge.

A second obstacle with which one has to deal is the absence of adequate means of measuring intelligence. Since there is no single test of intelligence against which serious objections have not been raised, it was thought that the most valid results would be secured by using a number of tests. The tests were given to the two groups, A and B. Group B will be discussed first as the results are more complete for it. There were 25 pupils present for all the tests given, the others being absent so much that the results were not considered. The following table shows the rankings of the pupils in each of the tests together with the final rank of each pupil in all the tests. The final rank was determined by adding together a pupil's ranks in all the tests except the vocabulary test. The smallest sum would then represent the highest rank.

TABLE X

Vocabulary	Knox Healy	Binet-Simon	Courtis Arith.	Starch Reading	Starch Spelling	Final Rank
1	17	1	5	4	1	2
2	8	3	1	1	14	1
3	23	8	7	2	3	5
4	14	2	12	7	2	4
5	5	6	8	9	5	3
6	6	23	23	13	10	18
7	15	15	21	10	4	13
8	21	7	15	11	9	12
9	22	10	4	6	8	7
10	4	9	10	5	16	6
11	18	4	14	17	6	9
12	12	13	16	20	23	21
13	20	22	21	3	24	23
14	16	5	8	8	13	8
15	13	14	3	24	7	11
16	11	11	10	23	19	16
17	19	19	18	25	21	25
18	10	12	19	16	11	14
19	1	20	19	15	20	18
20	25	24	2	22	17	22
21	9	17	13	19	12	15
22	3	16	16	21	18	16
23	2	18	6	12	22	10
24	24	25	23	14	15	20
25	7	21	25	18	25	24

The Knox-Healy* tests consist of a series of form boards, pictures, blocks, combination locks, jig saws, etc., designed to test the general intelligence of an individual as it functions in observation of form and details of a situation, in observation and memory of a series of movements, and in general manipulative and motor ability. The coefficients of rank correlation between the vocabulary test and the other tests were as follows: Binet-Simon, .66, Knox-Healy, −.11, Courtis Arithmetic, .36, Starch Reading, .67, Starch Spelling, .69. The coefficient for the vocabulary test and final rank is .69, and for average scholarship .59. The fact that there is a negative correlation between the Knox-Healy and the vocabulary test probably bears out the belief that there is little relation between verbal and motor ability. Naturally, the most significant correlation is that between final rank and vocabulary ability since the final rank in intelligence is based upon five, rather than one test.

Only 20 of the 25 in the second group were present for a sufficient

*This test was applied by Mr. Clifford Granger, a graduate student in Professor O'Shea's seminary in Education, and to him the author is indebted for the results.

number of the tests to make their results worth using, and a number of these were not present for all the tests.* It thus happens that the final ranks for a few had to be determined upon the basis of three or four different tests. The rank for each pupil in the Courtis tests is his rank in the four tests in the fundamentals. The "A" test was given once a week for a period of eight weeks and the rank is indicative of the rate of improvement during the period. The two composition tests were given to discover the range of words for each pupil in writing a fifteen minute composition on a familiar topic. In the association test twenty simple words were pronounced at intervals of one minute, the pupil writing down during the time all the words which came to his mind. His rank depends upon the number of different words which he wrote.

The coefficient of rank correlation between the vocabulary test and the final ranks in Table XI is .60. It is not felt that this coefficient is so significant as was the corresponding one in the former group since the tests are predominantly verbal tests.

TABLE XI

Vocabulary	Comprehension in reading	Speed and Comprehension in reading	Courtis Arithmetic	"A" test	First Composition	Second Composition	Association test	Final rank
1	1	2	..	2	4	1	..	1
2	..	..	17	..	..	8	4	12
3	2	4	4	1	7	5	5	3
4	15	11	18	5	..	..	8	16
5	5	7	8	..	5	6	5	5
6	10	12	15	..	..	..	3	13
7	5	3	2	7	..	2	2	2
8	4	1	13	11	..	17	1	7
9	2	8	3	..	6	3	..	4
10	7	5	6	..	11	7	12	9
11	..	..	10	3	8	10	9	8
12	12	10	1	4	1	14	7	6
13	16	15	9	..	..	13	..	20
14	13	16	6	10	12	9	11	14
15	..	..	15	9	10	16	10	17
16	8	13	10	6	2	11	14	11
17	19	13	..	8	3	12	13	15
18	..	..	4	..	9	4	15	10
19	14	9	12	12	13	15	16	19
20	10	5	13	..	14	18	..	18

*For the privilege of making the studies of this group, as well as for some of the material here used, the author is indebted to Professor V. A. C. Henmon.

It is now possible to make a sort of survey of the whole study and note some of the prominent indications. It has been shown that there is a large degree of correlation between a pupil's ability in oral expression and his vocabulary ability. It frequently happens that a child who is quiet and reserved does not get credit for a good vocabulary when he possesses one, while the child who talks freely is rated by the casual observer as a better linguist than the former though he may have a meager fund of words. In Group A, for example, the pupil ranking 19 in vocabulary is an incessant talker, but the number of words used in a given time, either in speaking or writing, is relatively small. Moreover, her expression, when carefully analyzed, is neither meaningful nor elegant. In Group B the pupil who ranks tenth in vocabulary is also of this type. The teacher experiences continual difficulty in keeping him from talking constantly in the recitation; while the pupil who ranks first in vocabulary is quiet, reserved and reflective. In speaking, the former is impulsive, inaccurate and inelegant while the latter is accurate, careful and discriminating.

With written expression the case is fairly similar. A number of regular school compositions from Group B were examined and, with very few exceptions, in which the deviations were not large, it was found that the pupils ranking high in vocabulary expressed their thoughts in more and better chosen words than did those whose vocabularies were below the average.

The main question with relation to the correspondence between language ability and average scholarship grades is, in what degree are average scholarship grades indicative of general intelligence? Some psychologists maintain that school performance is to no extent, or at least only to a very small extent, indicative of intelligence. Naturally, it cannot be affirmed that school grades are absolute and accurate measures of intelligence. The fact that practically all mental tests agree in indicating much greater individual differences than school grades show, is good evidence that school grades are not finely discriminating. This fact, however, does not mean that school activities do not test intelligence, nor that school grades are not, on the whole, indicative of mental capacity. School grades are to some extent determined by matters of expediency and economy in school administration and organization; but the statement so often made, that regular school activities are too narrow to test the various phases of the child's mental

capacity, lacks verification under present day school conditions. The facts already established by this investigation, of the high correlation between size of vocabulary and average school grades on the one hand, and between size of vocabulary and various mental tests on the other, tend to substantiate the assertion that average scholarship grades are indicative of general intelligence.

Now if this is established, it follows that the size of one's vocabulary under normal conditions is also a fairly accurate measure of his general mental capacity. This conclusion is supported by the results of studies of groups A and B, where it is found that, if several types of tests are given and the estimates of intelligence based upon the results, there is a good correlation between these results and verbal ability.

What, then may we conclude as to the exact relation between language and thought? We have referred to Preyer, who contends that the child shows evidence, through various activities, of extensive thought processes or deliberation before he acquires the use of language. As opposed to this view, it has been shown by Morgan (11) that animals often carry on activities which apparently imply just as complex thought processes as are necessarily connected with the cases cited by Preyer. Is it not then probable that whatever mental activities the child performs before he learns the use of language are precisely comparable to those which are common to the higher types of animals? Morgan says, "Language, and the analytical faculty it renders possible, differentiates man from the brute" (11, p. 374). Baldwin makes a similar statement when he says that with the getting of concepts "as opposed to the recepts of the animals . . . goes the development of speech, which some psychologists consider the source of all man's superiority over the animals" (1, p. 42). It cannot be doubted that animals have extensive processes of imagery and associations, but it is also true that there is a distinct difference between them and the mental processes of a rational human mind. Now this difference seems to be due to the capacity of the human mind for generalization, a capacity which is quite dependent upon symbols. And just to the extent to which these symbols make generalization possible, will this generalization be able to extend itself. Evidently then, the degree of intelligence which an ordinary person would manifest in consciously thinking out his reaction to any given stimulus, depends upon just how extensively the process of generalization has been

carried out. Without language, which is a highly refined system of symbolism, the generalizing process certainly cannot proceed. The child before learning to use words, and to some extent man throughout life, carries on the same sort of deliberation that the lower animals do, that of a very simple, direct, and concrete association. After he begins to learn words and, in fact, coincident with the word learning process, comes the conceptual activity. Morgan would even deny the child mind without language real percepts, preferring the term "mental products of a perceptual order." It is, however, quite certain that he does not account for the difference between child mind and brute mind which is apparent before the child begins to use language directly, that is, the manifest tendency of the child mind to generalization of experience, shown in anticipatory conduct. This is pointed out by Baldwin. This anticipatory conduct, which is the process or tendency of the mind in predicting results or acts through apperception, simply implies the indirect use of the language the child hears in organizing his own experience before he is able to make the speech organs function properly. This, then, leads to the conclusion that a "vague and confused" sort of generalization is taking place in the child mind before he uses language in a direct manner; whence it becomes perfectly clear that speech and thought are developing side by side in the child, each aiding and reenforcing the other. Just this lack of the indirect speech element is what seems to distinguish the brute mind from the human mind without language.

Now it seems perfectly rational, to repeat our former statement, that in any individual intelligence is conditioned by the extent to which he has carried out the generalizing of his experience, the building up of his conceptual order, which is, as we have just seen, almost wholly dependent upon the degree in which he is able to use highly refined symbolism. We may say, then, that language becomes a kind of filing case in which experience is sorted out, classified, and organized, and that without it, mental associations must remain of a primitive, direct, and simple type.

Pedagogical Aspects

From a number of investigations (6, p. 117; 2; 3) it has been determined that normal children from three to five years of age utter, under ordinary circumstances, from 1000 to 1400 words per hour during the entire day. In the author's study of child language

this has been verified both in the case of a three-year-old child and in that of one four and one-half years. In this particular study it was also found that the child of four and a half years, if allowed perfect freedom in talking, was not linguistically inactive for a period exceeding three minutes at any one time during the entire day. From the very nature of our present school system such freedom and continuity of oral expression is impossible, even if desirable. To discover in what degree children in the school room do have opportunity for oral expression and what the character of the expression is, the author spent 70 hours observing this phase of language work in the various grades. In a fourth grade, where 15 hours were spent, the highest number of words spoken per hour by pupils was 992, the lowest 416, and the average 705. The average number of pupils present during the time was 25. These words include words uttered in oral reading, although periods were chosen in which there was comparatively little of this. Ten periods were spent in kindergarten or primary grades with similar results as to number of words spoken; the remainder of the time was spent in the upper grades. In all grades under normal conditions the rate of speaking for hour periods seemed to be about the same as that given above. In the upper grades, naturally, there were longer periods of silence and the rate of speaking by pupils was at times much greater. It seems obvious that the number of words spoken by pupils cannot greatly exceed these estimates since the teacher generally talks at least one-half the time. Now if the number of words spoken per hour, as indicated here, be divided by 25, the number of pupils in the room, it is seen that the average for each pupil is, at most not over 40 words per hour. What happens is, of course, that a few pupils talk a great deal while several say little or nothing. The question at once arises, if there is an expressive instinct, or a natural tendency to expression which prompts the child to tell his experiences to his associates, to respond verbally to all sorts of stimuli continually, to what extent may the school without danger repress this tendency. Teachers often ask why pupils become timid at a certain period and reluctant to get up and speak or recite. The explanation appears to be found largely in this continual repression of the natural impulse to communication. The timidity does not appear suddenly or at any definite period but is the result of a gradual growth which is fostered and nourished from the day the child enters school. The child is reproved re-

peatedly because he speaks either to the teacher or to other pupils without permission and finally he often comes to associate a feeling of timidity with speaking in school. In many cases this association is built up to such an extent that there is a distinct feeling of unpleasantness, even of dread, upon the part of the pupil upon hearing his name called in recitation. Some repression is undoubtedly beneficial; pupils must learn to listen attentively as well as to speak accurately; they must learn that others have a right to be heard. Nevertheless, it can hardly be doubted that the process through which pupils are compelled to inhibit this social tendency should be a more gradual one and it is very doubtful if it should ever be carried so far as it is in the present school system.

BIBLIOGRAPHY

1. Baldwin, J. M. *The Story of the Mind.*
2. Bell, Sanford. Independent, 1903. Vol. 55, pp. 911–914.
3. Brandenburg, G. C. Ped. Sem. March, 1915. 89–120.
4. Chubb, Percival. *The Teaching of English.*
5. Dewey, J. *How We Think.*
6. Gale, I. M. C. & H. *Psych. Studies,* No. 1, July 1900.
7. Kirkpatrick, E. A. Pop. Sci. Mo. 1907, Vol. 70, pp. 157–164.
8. Knox, H. Scientific American, Jan. 9, 1915.
9. McMurry, Charles. *Special Method in Language.*
10. Meyer. *The Fundamental Laws of Human Behavior.*
11. Morgan, Lloyd. *Animal Life and Intelligence.*
12. Muller, F. M. *The Science of Language.*
13. Myers. *Experimental Psychology.* Parts I & II. (On Methods of Correlation)
14. O'Shea, M. V. *Linguistic Development and Education.*
15. Preyer, W. *Infant Mind.*
16. Schadel, E. *Das Sprechenlernen unserer Kinder.*
17. Stern, W. *The Psychological Methods of Testing Intelligence.*
18. Thorndike, E. L. *Mental and Social Measurements.*
19. Wundt, W. *Outlines of Psychology.*
20. Wundt, W. *Human and Animal Psychology.*

AN ABSOLUTE POINT SCALE FOR THE GROUP MEASUREMENT OF INTELLIGENCE.—PART II.

ARTHUR S. OTIS

Surgeon General's Office, Washington, D. C.

VII. Overlapping of Ability Between Grades

The points in Fig. 1 belonging to pupils in the eighth grade were made as circles, those belonging to pupils in the sixth grade, crosses, and those belonging to pupils in the fourth grade, triangles. It will be noted that there is considerable overlapping between the grades even though they are not consecutive.

The average score of the fourth graders is 385, of the sixth graders, 514; and of the eighth graders, 605. Suppose we call the norm for the third grade 320, the norm for the fifth grade 450, and for the seventh grade, 565, as shown in Fig. 1. We then find 8 fourth graders out of 43 above the fifth grade norm. Presumably these could do satisfactory fifth grade work. We find 1 of these 8 above the sixth grade norm. And we find 4 fourth graders below the third grade norm. The scattering of the three grades is shown in Table VI.

TABLE VI

Showing the Overlapping between the Grades

Norms:		3rd	4th	5th	6th	7th	8th
Fourth Grade (43)	4	19	12	7	1		
Sixth Grade (40)			5	11	15	4	5
Eighth Grade (38)				3	5	10	20

Another rather interesting fact concerning the distributions of scores, particularly in the sixth grade, is that there is a tendency for the more mature* pupils, intellectually, to be the younger ones. It would seem from this and the many other similar investigations that this is invariably the case. The most mature pupil in the sixth grade is, in fact, next to the youngest, while the oldest is next

*It has been necessary in this case to avoid the use of the ambiguous word, intelligence, which is used by nearly all writers on mental testing to mean both maturity, irrespective of age, and brightness—maturity with respect to age. The statement that, in a single grade, the youngest are also the most intelligent, according to the second meaning, would be a mere platitude. This would be true even if there were zero correlation between age and maturity.

to the least mature. In the eighth grade, also, the youngest is more mature, intellectually. There is, in other words, a negative correlation between age and maturity in the single grades. If pupils were graded according to intellectual maturity only, there would be no appreciable correlation, positive or negative, between age and maturity in a single grade. The act of negative correlation, therefore, suggests strongly that some bright pupils (mature but young) have been held back by the inelastic system of grading and that dull pupils have been promoted beyond their ability. This is one of the evils which mental testing should eventually remedy.

VIII. The Refinement of the Scale

Finding the Order of Difficulty of the Elements of the Tests. While not essential, it is nevertheless very desirable to have the elements of each test arranged in the order of difficulty. The relative degrees of difficulty of the elements of a test are, of course, probably not the same for any two individuals. The best arrangement, however, is probably the order of the elements according to the number of individuals who pass each, beginning, of course, with the easiest. In order to determine this ranking, the number of individuals who failed in each element was found during the scoring, for the Spelling, Arithmetic, Synonyms, Proverbs, and Relation Tests. To give an idea of the distribution of difficulties of the elements of these five tests in Scale 1, Fig. 3 was made. The horizontal position of each circle represents the number of individuals who failed in a given element. The circles at the left, therefore, represent easy elements. It is apparent from this as well as other sources that the Spelling Test is too easy for this group of individuals. The elements should be of such difficulty that the median element, in difficulty, is passed by about 50% of the group. The Synonym Test is somewhat too easy. The Arithmetic problems appear to fall into two distinct groups in difficulty. Problems of medium difficulty should be substituted for some of the others. The distributions of difficulty in the Relation and Proverb Tests are, perhaps, fairly satisfactory.

The Diagnostic Value of the Single Test Elements. It is not deemed within the scope of this study to investigate the value of each element of each test, as for example, a single problem in Arithmetic, as a measure of general ability such as is measured by the total point score of an individual, or of general arithmetical ability as measured

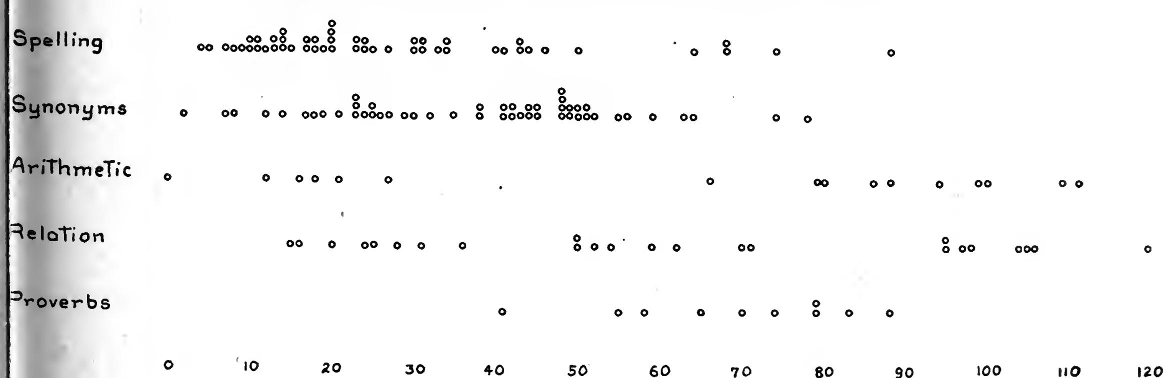
by the arithmetic score. However, as suggestive of means by which this may be done, we have examined the sixteen elements of Arithmetic Test I with the view to discovering which were the most suitable to be included in a test designed to be part of a scale for measuring general ability.

The method employed was as follows: The 121 individuals were first ranked in order of their total point scores. The papers of the 121 individuals in the Arithmetic Test I were then arranged in the same rank order. The sequence of passings and failings of Problem 1, Problem 2, etc., were noted. These are represented in Fig. 4 for each of the 16 problems and for the 121 individuals. There is one dotted line for each problem; each dotted line contains 121 units, one for each individual in the order of their total point scores. The presence of a unit of line indicates a problem correct; the absence of a unit line indicates a problem failed. The lines are arranged in the order of difficulty of the problems beginning with the easiest, the numbers of the problems represented are given at the left. The number of passes for each problem is represented by the position of a small circle on the line. In this figure the relative values of the problems as measures of general ability are shown by the relative amounts of overlapping of passes and failures—the greater the overlapping, the less the diagnostic value of the problem.

If the range of abilities of the individuals tested had been sufficiently broad so that the complete range of overlapping was represented for each problem, it would be a comparatively simple matter to express the relative diagnostic value of each problem by a single

FIG. 3

Showing the Distributions of Difficulties of the Elements of Five-Tests



number. For example, let us suppose that no individuals having measures of general ability above those represented in the figure would fail in either of Problems 7 and 8 and that no individuals having measures of general ability below those represented would solve either of these problems. If there were no overlapping in either case a full line would extend just to the circle representing the total number of passes, these being 94 and 55 respectively. Since in line 7 there are 10 failures before and 10 passes after the circle, we could represent the amount of overlapping by the number 10. Similarly since in line 8 there are 23 failures before and 23 passes after the circle, could represent the amount of overlapping in this case by the number 23. A rank order of the problems according to these numbers would, under the conditions mentioned above, give a serviceable indication of the comparative diagnostic values of the problems.

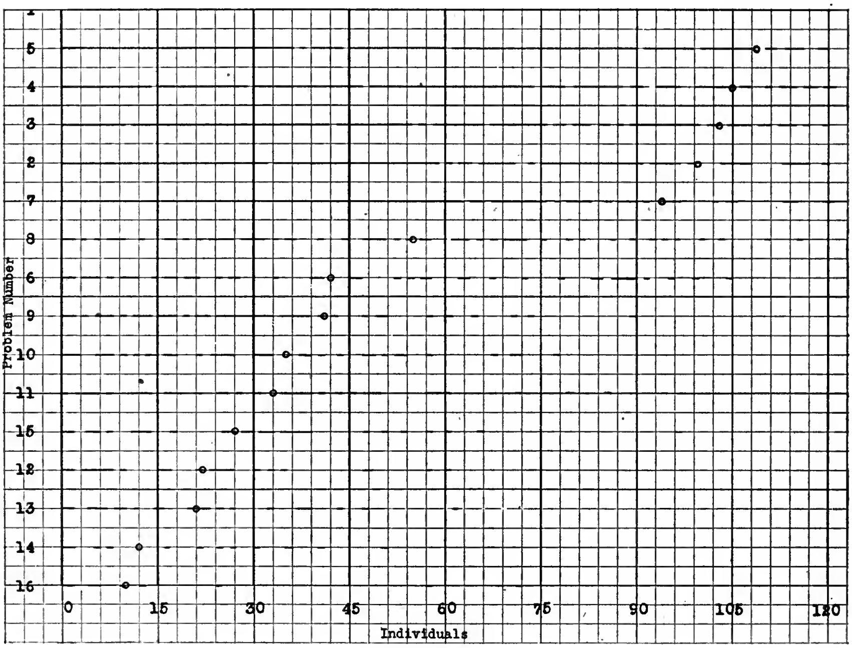


FIG. 4

Showing the Rank Order in Intelligence of the Individuals who Passed Each Problem of Arithmetic Test I.

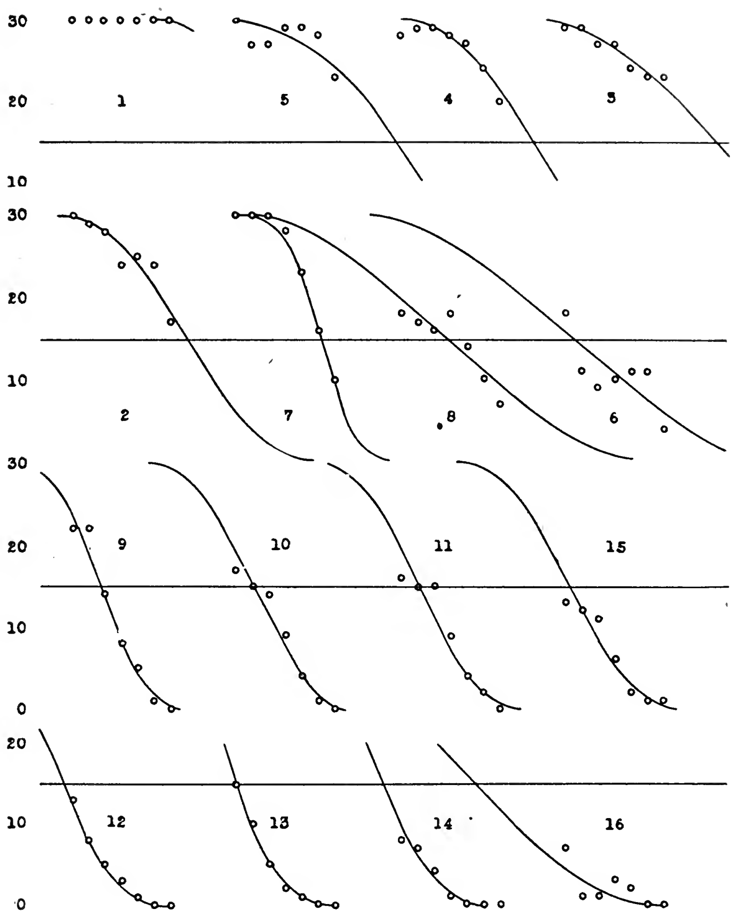


FIG. 5

Showing the Diagnostic Value of Each Problem of Arithmetic Test I

Inasmuch, however, as only a portion of the overlapping is represented in each case, it becomes necessary to adopt further means of ascertaining the relative diagnostic values. A method offering greater refinement is illustrated in Fig. 5. In this figure each group of seven circles pertains to one problem. The heights of the first circle, according to the scale at the left, represents the number of passes by the first 30 individuals in rank order. The height of the second circle represents the number of passes by individuals 16 to 45, inclusive, the third circle, individuals 31 to 60, etc., embracing 30 individuals in each group, the last group being 91 to 120. There is, of course, a tendency in each case for the succeeding numbers of passes to decrease. Theoretically, there should be a tendency for the circles to lie in a smooth curve of the form of an ogive. The steepness of the curve would indicate the degree of diagnostic value of the problem. The merit of this method is that it is possible in many cases to obtain a fairly good idea of the true slope of the curve for a given problem from only partial data. Curves have been drawn in what was judged by the eye to be approximately the true position of the curve. The problems may now be ranked in diagnostic value according to the slope of the curve at the 50% point. Thus it will be seen that the diagnostic value of Problem 8 appears to be the best. The value of Problem 1, of course, cannot be found. Quite possibly it would be as good as the average for lesser degrees of ability. For this group of individuals, of course, it has no value except perhaps for illustrated purposes. With the exception of Problems 6 and 7, possibly of 3 and 16, it would seem that the diagnostic values of the problems may be considered satisfactory.

It should be noted that the horizontal position of the point at which the curve crosses the 50% line affords a refined measure of the degree of general ability to which the ability to solve the problem corresponds. Thus, Problem 8, for example, may be considered "standard" for the degree of general ability slightly less than that of the 90th individual in rank order, say of the 93rd; or of 600 points. This method is practically that suggested for the standardization of tests of the Binet Scale (see Ref. 15). To thus express the degrees of difficulty of the several problems in terms of the Absolute Scale would assist in making the increments of difficulty between the problems equal. Such an amount of refinement, however, is not considered to be of great value until more of the obvious defects of the scale have been eliminated.

To determine the relative values of the problems as measures of "arithmetical ability," it would be necessary, of course, merely to rank the papers in the order of the arithmetic scores instead of the total scores.

IX Inter-Test Correlations

Considering the fact that double measures were obtained of each ability tested and that the number of individuals was comparatively large, it was deemed valuable to obtain the inter-test correlations.

These are given in Appendix IV. They are correlations between measures obtained in each case by combining the results of the two tests of a kind. There are given both the same coefficients and those corrected for attenuation due to errors of measurement. The formula used for correcting for attentuation was as follows (see Ref. 3).

$$r_{ab(corrected)} = \frac{r_{ab(raw)}}{\sqrt{r_{aa}\ r_{bb}}}$$

It is considered necessary to leave the discussion of the inter-correlations to a later article.

X. Further Considerations Regarding Reliability

The Reliability Coefficient of the Point Scale. To find the reliability coefficient of correlation between Scale I and Scale II, we may proceed as follows. Let us call the ten tests of Scale I $a_1, b_1, c_1, --- j_1$, and those of Scale II, $a_2, b_2, c_2, --- j_2$. Then by the formula (See Ref. 13) for the correlation of the sums of several variables, the standard deviations of the distributions of scores in the several tests having been made equal,

$$r_{(a_1+b_1+c_1+--+j_1)(a_2+b_2+c_2+--+j_2)} =$$

$$\frac{r_{a_1a_2}+r_{a_1b_2}+r_{a_1c_2}+\ldots+r_{b_1a_2}+r_{b_1b_2}+r_{b_1c_2}+\ldots+r_{c_1a_2}+r_{c_1b_2}+r_{c_1c_2}+\ldots}{\sqrt{10+2(r_{a_1b_2}+r_{a_1c_1}+\ldots+r_{b_1c_1}+\ldots)}\ \sqrt{10+2(r_{a_2b_2}+r_{a_2c_2}+\ldots+r_{b_2c_2}+\ldots)}} \quad (1)$$

$$= \frac{(\text{Sum of 10 reliability coefficients, } r_{a_2a_2}\text{, etc.})+2(\text{Sum of 45 coef. of intercorrelation, } r_{a_1b_2}\text{, etc.})}{\sqrt{10+2(\text{Sum of 45 coefs. of intercor., } r_{a_1b_1}\text{, etc.})}\ \sqrt{10+2(\text{Sum of 45 coefs. of intercor., } r_{a_2b_2}\text{, etc.})}}$$

But since the correlations, $r_{a_1b_1}$, $r_{a_1b_2}$, and $r_{a_2b_2}$, tend to be equal, we may take the sums of each of the three sets of 45 coefficients of

correlation to be equal to one another. We may therefore simplify the equation thus:

$$r_{(a_1+b_1+c_1+\ldots_1+j_1)\ (a_2+b_2+c_2+\ldots+j_2)} = \frac{\Sigma r_{xx}+2\Sigma r_{xy}}{10+2\Sigma r_{xy}} \qquad (2)$$

in which $r_{xy}=r_{aa}+r_{bb}+r_{cc}+\ldots$
and $r_{xx}=r_{ab}+r_{ac}+r_{ad}+\ldots+r_{bc}+\ldots$

Since the intercorrelations were found between *double* measures in each test, it is necessary to express the reliability coefficients also in terms of double measures. As these were found in terms of single measures, each has been transmuted into terms of double measures by means of the formula,*

$$r_{2a2a}=\frac{2r_{aa}}{1+r_{aa}} \qquad (3)$$

in which r_{2a2a} and r_{aa} are respectively the correlations between double measures and between single measures of any ability. The reliability coefficients in each test in terms of double measures are shown in the fourth column of Table III. Their sum (Σr_{xx}) is shown to be 8.877. The sum of the 45 coefficients of intercorrelation multiplied by $2=56.478$ ($=2\Sigma r_{xy}$). Solving formula 2 above,

$$r_{(\text{Scale I}_2)\ (\text{Scale II}_2)}\ \frac{8.877+56.478}{10+56.478}=.983.$$

This is when Scale I_2 and Scale II_2 are considered as double scales. To find the reliability coefficients of correlation between Scale I and Scale II as single scales, *i. e.*, each composed of the tests truly comprising it (call these Scale I* and Scale II*), then, according to the formula,†

$$r_{aa}=\frac{r_{2a2a}}{2-r_{2a2a}}$$

in which r_{aa} and r_{2a2a} have the same meanings respectively as before,

$$r_{(\text{Scale I}_1)\ (\text{Scale II}_1)}=\frac{.983}{2-.983}=.967$$

This, then, is the reliability coefficient of the Point Scale.

The Probable Error of the Point Scale

As has been shown (see Ref. 11)

$$r=1-\frac{\sigma_e^2}{\sigma_{dist.}^2} \qquad (1)$$

*This formula is a corrollary to formula 1 above.

†The inverse of formula 3.

in which r is the reliability coefficient of correlation between two series made up of pairs of measures, a_1, a_2; b_1, b_2; c_1, c_2; etc.; in which σ_e is the standard deviation of the errors of measurement of a, b, c, etc.; and in which $\sigma_{dist.}$ is the standard deviation of the distribution of values, a, b, c, etc., in either series.

From equation 1 it follows that

$$\frac{\sigma_e^2}{\sigma_{dist.}^2} = 1 - r \text{ whence } \sigma_e^2 = (1 - r)\,\sigma_{dist.}^2$$

$$\sigma_e = \sqrt{1 - r}\,\sigma_{dist.} \qquad \text{and} \qquad \text{P. E.} = .6745\sqrt{1 - r}\,\sigma_{dist.} \qquad (2)$$

in which *P. E.* is the probable or median error of measurement.

If, now, we consider r to be the reliability coefficient of the Point Scale, P. E. as the probable error of measurement by either Scale I or Scale II, and $\sigma_{dist.}$ as the standard deviation of the point scores by the same scale, then we may solve equation 2 for the probable error of the scale. The standard deviation of the distribution of scores by the scale (average of both) was found to be 111 points. Solving equation 2,

$$\text{P. E.} = .6745\sqrt{1 - .968} \times 111 = 13.7$$

The Probable Error of the Point Scale, therefore, is 13.7 points.

This is 2.7% of the median score (500) of the whole group, or 3.0% of the total range of scores (461).

To view the reliability of the scale from another angle we may determine as nearly as possible the probable error of a mental age by the scale. Thus, if we may assume that the reliability co-efficient of correlation between mental ages by the scale is approximately equal to the reliability coefficient of correlation between the point scores and that the distribution of mental ages by the scale is approximately equal to the distribution of ages, then we may let P. E., in formula 2 above, represent the probable error of a mental age by the scale, r represent the reliability coefficient of correlation between mental ages by the two scales, and $\sigma_{dist.}$ represent the standard deviation of the distribution of mental ages. Then substituting in equation 2 the approximate values of these quantities, the standard deviation of the distribution of ages being 28.1 months, we have

$$\text{P. E.} = .6745\sqrt{1 - .967} \times 28.1 = 3.44$$

The Probable Error of a mental age by the Point Scale may be considered, therefore, as approximately 3½ months. This is practically the same as the probable error of a mental age by the Stanford Revision of the Binet Scale (see Ref. 11).

XI. Comparisons with School Mark and Amount of Schooling

Correlation of Total Score with School Mark. The teacher of each of the grades furnished for each pupil a final mark representing the relative character of his school work for the year. As the marks given by the different teachers were not comparable, a separate correlation was made for each grade between the school mark and the total point score. The coefficients were found for the fourth, sixth, and eighth grades to be respectively .80, .41, and .50. Since we have no measures of the reliability of the teachers' marks we are unable to determine the probable true correlation between intelligence, as measured by the scale, and school performance.

The Relation of the Coefficient of Brightness to the Amount of Schooling. The pupils were asked to tell the length of time they had spent in each grade. From these data, the total amounts of time spent in school was found for each pupil. For convenience, the 121 pupils were then classed together in groups according to the amount of retardation or advance. The number of the pupils of each class are shown in Table VII. Here again we have no measure of the reliability of the reports of schooling and are therefore unable to determine the value of the results. There is, however, a very definite tendency for advanced pupils to obtain high coefficients of brightness and for retarded pupils to obtain low coefficients of brightness.

TABLE VII

Showing the Amounts of Retardation and Advance and their Relation to the Coefficients of Brightness. (Data from 10 pupils were missing.)

	Retarded			At	Advanced	
	3 yrs.	2 yrs.	1 yr.	Grade	1 yr.	2 yrs.
Number of Pupils	2	8	33	56	8	4
Average Coef. Brightness	77	88	96	102	103	114

Appendix III.

Some Mathematical Reasoning with Regard to Criteria of Tests of Intelligence.

1. If we were to assume that each test measured only a general factor—one common to all the tests—and one or more factors specific to that test alone, then the relative degrees in which two tests correlate with the general factor, are expressed, subject to the

chance errors of the coefficients, by the relative degrees to which these two tests correlate with the other tests. This may be shown as follows. By formula for partial correlation,

$$r_{ac.i} = \frac{r_{ac} - r_{ai}\, r_{ci}}{\sqrt{(1 - r^2_{ai})\,(1 - r^2_{ci})}}$$

in which *a* and *c* are tests, *i* is a hypothetical, perfect measure of the general factor, r_{ac} is the coefficent of correlation between *a* and *c*, etc., and $r_{ac.i}$ is the coefficient of correlation between *a* and *c* which is due to factors other than the general one. But since by hypothesis, the general factor in the only source of correlation, $r_{ac.i} = 0$.

Then $\qquad r_{ac} = r_{ai}\, r_{ci}$

and similarly, $\qquad r_{bc} = r_{bi}\, r_{ci}$

Therefore $\qquad \dfrac{r_{ac}}{r_{bc}} = \dfrac{r_{ai}}{r_{bi}}$

Similarly, $\qquad \dfrac{r_{ai}}{r_{bi}} = \dfrac{r_{ad}}{r_{bd}} = \dfrac{r_{ae}}{r_{be}} = \dfrac{r_{af}}{r_{bf}} = \text{etc.}$

An expression for the combined value of these ratios is given very approximately by the ratio of the average intercorrelation between *a* and the other tests to the average intercorrelation between *b* and the other tests.

2. Let us consider now a case in which there is no factor common to all the tests in the group. To take a very simple example, let us suppose we have four tests (nos. 1, 2, 3, and 4) testing abilities each of which is made up of five of the nine elements, A, B, C, D, E, F, G, H, and I, distributed as follows.

Test 1,	A	B	C	D	E				
Test 2,	A	B	C	D		F			
Test 3,			C	D	E	F	G		
Test 4,		B			E	F		H	I

Here it will be noted that no element is common to more than three abilities. Now the coefficient of correlation between two series of values is a measure of the percentage of elemental causes common to both.* And since the number of elemental causes common to

*For example, if five coins are tossed *n* times and each time the number of heads appearing is recorded, and if after each independent tossing, one coin is left lying, the other four tossed again, and the number of heads then appearing is recorded; then as *n* approaches infinity, the coefficient of correlation between the number of heads appearing by the independent tossing and the number of heads appearing by the dependent tossing approaches .20, attesting to the fact that one fifth of the causes affecting the number of heads in each throw (one coin in five) was common to both throws of a pair. If two of the five coins are left lying, the correlation will approach .40, if three are left lying, .60, if four, .80, and if five, of course, 1.00. Similarly for other numbers of coins and similarly for elements of abilities.

abilities, 1 and 2, is four out of five, the correlation between tests, 1 and 2, will tend to be .80. Three elements are common to abilities, 1 and 3, and 2 and 3. Therefore the correlations between tests 1 and 3 and between 2 and 3 will tend to be .60. And so on. The correlation table will therefore appear as follows.

Tests	1	2	3	4
1		.80	.60	.40
2	.80		.60	.40
3	.60	.60		.40
4	.40	.40	.40	
Sums	1.80	1.80	1.60	1.20

A table showing the number of elements common to each pair of abilities would appear as follows.

Tests	1	2	3	4
1		4	3	2
2	4		3	2
3	3	3		2
4	2	2	2	
Sums	9	9	8	6

It may be seen from this table that the number of times the elements of ability, 1, appear in the other three abilities is 9. The correlation spread of ability, 1, may therefore be said to be represented by the number 9. The number of times the elements of abilities, 2, 3, and 4, appear in the other three abilities are respectively 9, 8, and 6. We may say, then, that the relative values of the correlational spreads of the four tests are as 9 : 9 : 8 : 6. Now it may be seen that these are exactly the same proportions as 1.80 : 1.80 : 1.60 : 1.20, the sums of the coefficients in the first table. The latter values are equal respectively to the former values when each is divided by 5, the number of elements in each ability. Thus it may be seen that the sums of the correlations of each of the tests with all of the others afford meaures of the relative correlational spread of the tests.

3. The coefficient of correlation between any one weighted test and the weighted composite of a number of tests may be found

from the coefficients of intercorrelation and the weights by the formula (see Ref. 13) which may be stated in general as follows.

$$r_{wa(x_1b_1+x_2b_2+x_3b_3\ldots)} = \frac{\Sigma({}_x\sigma_b r_{ab})}{\sqrt{\Sigma x\,\sigma_b^2+\Sigma xx\sigma_b\,\sigma_b\,r_{bb}}}$$

in which w_1, x_2, x, etc., are the weights given to the tests, a, b_1, b_2, etc., and σ_b is the standard deviation of the scores of any test, *b*.

McCall's procedure, therefore, might have been to consider b_1, b_2, etc., as representing the tests which he wished to embody in his Composite; to consider x_1, x_2, etc., as representing the respective weights to be given these tests, and *a* as representing any test it was desired to correlate with the Composite. The correlation could then have been obtained by solving the equation. The general formula, is equally applicable, of course, for finding, from the intercorrelations, the correlation of a test with the average of all the *other* tests.

If only the relative values of the correlations of each test with a composite of weighted tests is desired, these may be obtained more simply yet; thus, assuming that there were only three tests, *a*, *b*, and *c*, in the group, weighted respectively, *w*, *x*, and *y*, then the corresponding formula for the correlation of test *a*, with the weighted composite would be

$$r_{wa(wa+xb+y\ )} = \frac{w\Sigma_a r_{aa}+x\Sigma_b r_{ab}+y\Sigma_c r_{ac}}{\sqrt{w^2\sigma_a^2+x^2\sigma_b^2+y^2\sigma_c^2+2(w\sigma_a x\sigma_b r_{ab}+w\sigma_a y\sigma_c r_{ac}+x\Sigma_b y\Sigma_c r_{bc})}}$$

Similarly, $r_{xb(wa+xb+yc)} = \dfrac{w\Sigma_a r_{ab}+x\Sigma_b r_{bb}+y\Sigma_c r_{bc}}{\text{same denominator}}$

And $r_{yc(wa+xb+yc)} = \dfrac{w\Sigma_a r_{ac}+x\Sigma_b r_{bc}+y\Sigma_c r_{cc}}{\text{same denominator}}$

Since the denominators are the same in all cases, it may be seen that the relative values of the correlations of the several tests with the weighted composite are directly proportional to the sums of the intercorrelations of those tests, each with all the tests, when these intercorrelations have been weighted as shown in the numerators. And if the standard deviations have been made equal, this merely means weighting the several coefficients by the same weights in and the same order as they would appear in the composite. The same reasoning holds for any number of tests.

4. If it is desired, on the other hand, to find the relative or absolute amounts of the average intercorrelations of each of a series

with all the others, (weights and standard deviations being equal) as a criterion of the degree to which each test measures the common factor; and if the values of the separate intercorrelations are not required; it will be more convenient to derive these average intercorrelations from the correlation of each test with the average of all the measures taken together as a composite. That this may be done is shown as follows.

Repeating the proof given in 3 above in a simpler form, let us assume again for the moment that there are only three tests, *a*, *b*, and *c*, in the series; then by the formula for the correlation of one test with the average of a number of tests (assuming weights and standard deviations equal),

$$(1) \qquad r_{a(a+b+c)} = \frac{r_{aa}+r_{ab}+r_{ac}}{\sqrt{3+2(r_{ab}+r_{ac}+r_{bc})}}$$

$$(2) \qquad r_{b(a+b+c)} = \frac{r_{ba}+r_{bb}+r_{bc}}{\sqrt{3+2(r_{ab}+r_{ac}+r_{bc})}}$$

$$(3) \qquad r_{c(a+b+c)} = \frac{r_{ca}+r_{cb}+r_{cc}}{\sqrt{3+2(r_{ab}+r_{ac}+r_{bc})}}$$

Letting $\Sigma r_{a,b,c(a+b+c)}$ represent $r_{a(a+b+c)}+r_{b(a+b+c)}+r_{c(a+b+c)}$, and since $r_{aa}=r_{bb}=r_{cc}=0$,

$$(4) \qquad \Sigma r_{a,b,c(a+b+c(} = \frac{3+2(r_{ab}+r_{ac}+r_{bc})}{\sqrt{3+2(r_{ab}+r_{ac}+r_{bc})}}$$

$$(5) \qquad \Sigma r_{a,b,c(a+b+c)} = \sqrt{3+2(r_{ab}+r_{ac}+r_{bc})}$$

Multiplying equation 1 by equation 5, we have

$$r_{aa}+r_{ab}+r_{ac} = r_{a(a+b+c)} \times \Sigma r_{a,b,c(a+b+c)}$$

and similarly for all other correlational sums. Thus it may be seen that the *absolute* amounts of the sums or averages of the intercorrelations of any test with all the tests in the series may be derived from the values of the correlations of each test with the composite (weights and standard deviations being equal) without the individual test intercorrelations being found. The same reasoning holds for any number of tests. As a criteria of the degree to which any test measures the factor common to the group of

abilities tested, the test's average intercorrelation with all the tests and its correlation with the composite are of equal value, being, in fact, the same criterion. The sums of the intercorrelations of any test with all the *other* tests (excluding itself) may be obtained, of course, by subtracting 1.00 (the correlation of the test with itself) from the sum of the intercorrelations with *all* the tests.

REFERENCES

1. A. R. Abelson. *Mental Ability of Backward Children.* Brit. Jr. of Psy., Dec., 1911, 268–314.
2. Leonard P. Ayres. *Measurement of Ability in Spelling.* Russell Sage Foundation, Educational Monograph E139.
3. Cyril Burt. *Experimental Tests of General Intelligence.* Brit. Jr. Psy., Vol. iii, 94.
4. F. J. Kelly. *The Kansas Silent Reading Tests.* Bureau of Educational Measurements and Standards, Emporia, Kansas.
5. Wm. A. McCall. *Correlation of Some Educational and Psychological Measurements.* Columbia University Contributions to Education, No. 79.
6. M. Merriman. *Method of Least Squares.* John Wiley and Sons, N. Y., 1894, Ch. IV.
7. Arthur S. Otis. *The Reliability of Spelling Scales.* School and Society, Vol. IV, Nos. 96, 97, 98, and 99.
8. Arthur S. Otis. *Considerations Concerning the Making of a Scale for the Measurement of Reading Ability.* Ped. Sem., Dec., 1916, Vol. XXIII, pp. 528–549.
9. Arthur S. Otis. *Some Logical Aspects of the Binet Scale.* Psy. Rev., Vol. XXIII, Nos. 2 and 3.
10. Arthur S. Otis. *A Criticism of the Yerkes-Bridges Point Scale, with Alternative Suggestions.* Jr. Ed. Psy., Mar., 1917.
11. Arthur S. Otis. *The Reliability of the Binet Scale and Pedagogical Scales.* (To be published).
12. Daniel Starch. *Educational Measurements.* The Macmillan Company, 1916.
13. C. Spearman. *Correlations of Sums and Differences.* Brit. Jr. Psy., Vol. 5, 419–426.
14. Lewis M. Terman. *The Measurement of Intelligence.* Houghton Mifflin Co., 1916.
15. M. R. Trabue. *Completion Test Language Scales.* Col. Univ. Cont. to Ed., No. 77, 1916.

APPENDIX IV

Inter-test Correlations (Raw)

	Synonyms	Narrative Completion	Following Directions	Proverbs	Relation	Arithmetic	Spelling	Geometric	Disarranged Sentences	Digit
Synonyms		.798	.781	.828	.750	.722	.753	.674	.699	.322
Narrative Completion	.798		.815	.738	.742	.719	.787	.689	.654	.341
Following Directions	.781	.815		.762	.749	.686	.662	.716	.561	.420
Proverbs	.828	.738	.761		.738	.700	.686	.677	.567	.358
Relation	.750	.742	.749	.738		.746	.688	.719	.612	.282
Arithmetic	.722	.719	.686	.700	.746		.662	.687	.705	.324
Spelling	.753	.787	.662	.686	.688	.662		.632	.679	.259
Geometric	.674	.689	.716	.677	.719	.687	.632		.570	.301
Disarranged Sentences	.699	.654	.561	.567	.612	.705	.679	.570		.280
Digits	.322	.341	.420	.358	.282	.324	.259	.301	.280	
Average	.703	.698	.684	.673	.670	.661	.645	.630	.592	.321

Inter-test Correlations (Corrected for Attenuation)

	Synonyms	Narrative Completion	Relation	Proverbs	Following Directions	Arithmetic	Geometric	Spelling	Disarranged Sentences	Digit
Synonyms		.901	.881	.960	.887	.808	.769	.825	.819	.378
Narrative Completion	.901		.845	.830	.898	.779	.762	.836	.742	.386
Relation	.881	.845		.864	.858	.842	.828	.760	.723	.332
Proverbs	.960	.830	.864		.862	.780	.770	.749	.662	.416
Following Directions	.887	.898	.858	.862		.749	.798	.708	.642	.478
Arithmetic	.808	.779	.842	.780	.749		.753	.697	.793	.363
Geometric	.769	.762	.828	.770	.798	.753		.679	.655	.344
Spelling	.825	.836	.760	.749	.708	.697	.679		.748	.280
Disarranged Sentences	.819	.742	.723	.662	.642	.793	.655	.748		.329
Digits	.378	.386	.332	.416	.478	.363	.344	.280	.329	
Averages	.803	.775	.770	.766	.764	.729	.706	.698	.679	.367

The Journal of Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD.

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

NOTICE TO SUBSCRIBERS. Notify the publishers, Warwick & York, Inc., of any change of address; as copies lost in transit because of subscriber's removal to another address will not be duplicated.

THE JOURNAL is issuing a series of Monograph Supplements for the publicaiion of more extensive researches in educational psychology. For details prospective contributors may apply to H. E. Buchholz, 10 East Center St., Baltimore, Md.

The price of the JOURNAL is $3.00 per year of ten monthly issues, if paid in advance. Single issues are 50 cents each. Back volumes are $3.50 in original wrappers and $4.75 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME IX JUNE, 1918 NUMBER 6

EDITORIAL

Ninth Grade Mathematics

Between the vigorous attacks of practical business men and the critical skepticism of educational theorists the advocates of the traditional courses in mathematics are having a hard time. The business men complain that the graduates of our schools fail to meet even the simplest mathematical situations in commercial and industrial life, while the scientific students of education are confirming by the results of standard tests the lack of mastery which pupils show even of those aspects of the subject which they are supposed to have been taught. The teachers of mathematics feel themselves on the defensive, and are making valiant efforts to convince an increasingly doubtful public that mathematics is the *sine qua non* of a liberal education. Particularly strenuous are their exertions to maintain the tradition that mathematics is the best of all subjects "to train the reasoning powers," and they are not at all kindly disposed to the educational psychologists for raising the question as to what that expression means, or whether it has any meaning. Incidentally, it is vastly diverting to note their efforts to "prove" the value of mathematics

by industriously collecting the opinions of eminent divines, lawyers, physicians, and politicians as to the school subjects that have helped them most in life. If this is the kind of reasoning which their mathematical training has developed, then the seeker after truth may well ask to be delivered from it.

By all odds the sanest, most thorough, most helpfully constructive discussion of high school mathematics is to be found in Professor Rugg's recent monograph on "The Reconstruction of Ninth Grade Mathematics." Professor Rugg is himself something of a specialist in higher mathematics, is the author of a valuable treatise on educational statistics, and is wholly sympathetic in his attitude toward the study. Yet, on the basis of a most admirable survey of the history and present status of the subject, and after a very extended application of standardized tests, he states that "the candid judgment, joint and several, of teachers of advanced mathematical and scientific subjects, of employers of labor, and of educational critics and reformers agrees with the conclusion of the scientific investigator who has measured results more objectively: ninth-grade mathematics cannot merely be rearranged or "reorganized." *It must be rebuilt from the ground up.*"

Professor Rugg believes that with a complete reconstruction of the ideals, the content, and the methods of instruction of the subject in accordance with present social needs and in accordance with the learning processes of children, ninth grade mathematics *may* be made to contribute valuable training in "logical thinking." Very good. So may every school subject. Two questions are yet unanswered. Is this the *kind* of "logical thinking" that is involved in the "analysis, comparison, and discrimination" demanded by the problems of practical life? Does the study of such reconstructed mathematics furnish a *higher degree* of training in the "logical thinking" involved in those problems than would such subjects as economics, politics, sociology, or commercial geography, which deal directly with the data involved in that "logical thinking"? Obviously anyone who expects to do advanced work in mathematics or in any quantitative science must have a certain amount and a certain kind of elementary mathematics. Just how much and what kind need to be more carefully determined. But in view of the very small percentage of high school students who will ever engage in any such advanced work, it is extremely questionable whether even the reconstructed mathematics advocated by Professor Rugg should be required of every ninth-grade pupil.

J. C. B.

NOTES AND NEWS

The plan for promoting more intimate relations between France and this country by offering one hundred scholarships in colleges and universities to French women is meeting with great success. The French minister of public information has formally accepted the offer, and has stated that many applications have already been made by French women. Sixty-five American institutions have already offered such scholarships, usually two each, so that opportunities will be offered for more than the original one hundred planned for. A committee of women has already gone to France to make selections and assignments, and a reception committee has been organized to meet the students when they arrive in this country and assist them in getting comfortably settled in their respective institutions.

It is stated in the press that the Boston school committee has decided that it could not allow teachers who are now eligible for a sabbatical year, and who desire to spend it in the service of the Red Cross in Europe, the half-pay which they would customarily receive if they spent their sabbatical leave in "travel and study." Perhaps the school committee has some good reason for this action, but to the casual reader it seems singularly unpatriotic, inhumanitarian and short-sighted. It would be more in accord with patriotism for the school committee to collect all the teachers that could possibly be spared and send them to France on full salary and expenses. There is certainly no more humanitarian cause at the present time than to provide the best possible care and nursing for our boys at the front. And for purely selfish ends it is hard to see how a city could get a greater return for its money than to have as many of its teachers as possible come into first hand contact with conditions in France.

There has recently appeared from the press of Houghton Mifflin Company *The Life of Naomi Norsworthy*, by Frances Caldwell Higgins, with an Introduction by Dean Russell, of Teachers College. Those who ever came in contact with Dr. Norsworthy will appreciate the author's tribute to her "genius for friendship, her brilliant intellectual gifts, and her rare power as a teacher."

It is with regret that we note the name of Professor G. W. A. Luckey, dean of the graduate school of education at the University of Nebraska, among those professors whose resignations have been requested by the board of regents on account of alleged indiscreet and disloyal utterances.

The board of trustees of the University of Illinois have authorized the establishment of a college of education and a bureau of educational research. Dr. B. R. Buckingham, statistician of the Wisconsin State Department of Education, has been appointed director of the bureau.

Superintendent Ernest C. Witham, of Southington, Conn., is giving courses in educational measurements at the summer session of the New Hampshire State Normal School, Keene, N. H.

At Northwestern University Dr. Robert H. Gault has been promoted from associate professor to be full professor of psychology.

At the University of Chicago J. Franklin Bobbitt has been promoted from associate professor to professor, and William Scott Gray and Harold Ordway Rugg from assistant professor to associate professor in the department of education.

Dr. J. H. Coursault has been appointed dean of the school of education at the University of Missouri.

Dr. C. B. Cornell, of the University of Nebraska, has been appointed assistant professor of education at the University of Kentucky.

Professor W. E. Chancellor, head of the department of social sciences, College of Wooster, will teach educational sociology and school hygiene at the summer session of the new School of Education, Cleveland, which has been formed by a combination of departments from Western Reserve University and the City Normal School of Cleveland.

Professor John Dewey, of Columbia University, has been granted a year's leave of absence, beginning in September, and will deliver a series of lectures at the University of California.

Professor J. Mark Baldwin has been appointed lecturer in the Ecole des Hautes Etudes Sociales at Paris.

Dr. Florian Cajori, dean of the department of engineering and professor of mathematics at Colorado College, has accepted an appointment as professor of the history of mathematics in the University of California.

Dr. John J. Tigert, head of the department of psychology of the University of Kentucky, has obtained a leave of absence to engage in army Y. M. C. A. work in France.

Dr. Leonard P. Ayres, who has had charge of the statistical work of the War Department at Washington, has been made a colonel and is attached to the staff of General Pershing in France.

Superintendent E. S. Kuhnes, of Langdon, North Dakota, has entered the psychological clinic of the war service.

PUBLICATIONS RECEIVED

FELIX ADLER. *An Ethical Philosophy of Life.* New York: D. Appleton and Company, 1918. Pp. viii, 380. $3.00.

In this noteworthy book one of the leading American writers on ethical and social questions gives an account of the development of his moral philosophy. In Book I we learn why the author found the orthodox Hebrew religion unsatisfactory, what a strong appeal the doctrines of Emerson made, how profoundly he was influenced by the teachings of Jesus, and how he worked out his own doctrine of ethical culture and social reform. Book II gives us his reaction to Kant's philosophy upon which his own ethical system is based. Books III and IV deal with the applications of his theory to "the three great shadows, sin, sickness and death," and to the relationships of the family, of vocational activities, of the state, and of the several states to each other. The final chapter is entitled "The Last Outlook on Life," and gives the author's final summing up of the whole problem. As an epitome of a long and useful intellectual career, and as an account of a sincere effort to reach a basis for religion which would commend itself to the open-minded scientific inquirer, the book deserves careful attention.

CHARLES A. BEARD AND WILLIAM C. BAGLEY. *The History of the American People.* New York: The Macmillan Company, 1918. Pp. xv, 674. $1.20.

The youth of the land are indeed to be envied for the superior quality of the text books in elementary history that are being offered them. The present work has several points of distinction. The dominating motive of the book is "the preparation of children for citizenship through an understanding of the ideals, institutions, achievements, and problems of our country." "In carrying out our ideal we have selected those striking features of American history which bear upon and help to explain our own age. We have conceived of the whole as a vital, moving story with certain very definite and fundamental acts and scenes. We have sought to give to the book that unity which comes from such a controlling purpose, and have subordinated to it all details and collateral matter." Almost five hundred pages are devoted to the period since the close of the Revolution. There are two chapters on American education, several on industrial development, and particular attention is paid to the growth of democracy. The book is provided with four colored plates, twelve colored maps, and a large number of interesting illustrations. The story is brought down to December, 1917, to the government control of the railways.

CARL C. BRIGHAM. *Two Studies in Mental Tests. I. Variable Factors in the Binet Tests. II. The Diagnostic Value of Some Mental Tests.* Psychological Monographs, Vol. 24, No. 1, Whole Number, 102. 1917. Pp. 254. $2.50.

In the first of these studies we have a detailed examination of the individual tests of the Binet scale, on the basis of the responses of 422 elementary school pupils. The results are discussed with particular reference to (1) the personal equation of the examiner; (2) the correlation of each test with the standing of the pupils in the grades; and (3) sex differences. In some of the tests a very considerable difference was found in the results of different examiners, and the author joins Doll, Whipple, Kuhlmann and others in warning against the indiscriminate use of the tests by untrained and inexpert examiners. The tests showed a general correlation with school grade, but gave only slight evidence of sex differences. The second study

presents the results of the examination of another group of 229 boys with twenty-three of the Binet tests and a number of supplementary tests, from the point of view of the diagnostic value of each individual test. These data furnish a valuable basis for determining the significance of various tests. The author accepts the point scale principle as the most satisfactory, but objects to the point weighting given to various tests by Yerkes and others.

WALTER M. GALLICHAN. *The Psychology of Marriage.* New York: Frederick A. Stokes Company, 1918. Pp. xi, 300. $1.50.

Partly as a result of recent studies in social biology and partly due to the revelations of psychopathology, there is an increasing appreciation of the importance of sex relationships for normal, wholesome living. It is recognized by all thoughtful students of the subject that much marital trouble is due to ignorance, and in the present book an attempt is made to set forth in simple, direct narrative the essential findings of medical, social and psychological investigations on the subject of sex. It is not the traditional volume of trite advice to newly-weds, but a serious sociological study designed to stimulate reflection on the significance of marriage for the individual and the community, and to contribute to the development of a rational sex ethics that will meet the conditions of modern life.

MARY HARGRAVE. *The Earlier French Musicians.* New York: E. P. Dutton and Company, 1918. Pp. 258. $1.50.

With the subsidence of German domination in music there is a natural increase of interest in the music of other countries, particularly of France. This little volume on the development of French music to the beginning of the nineteenth century will, therefore, be quite opportune. We find an account of Lully, the popular opera and ballet composer and favorite of Louis the Fourteenth; Rameau and the early clavecine composers, the forerunners of modern piano music; and the group of operatic composers that flourished during the Revolution and the Napoleonic period, Gretry, Mehul, Lesueur, Cherubini, and Boieldieu. Charming and entertaining as many of these men and their works are, they are worlds removed from the art aspirations of such modern French musicians as Berlioz, Bizet, Saint-Saens, Cesar Franck, Debussy and Ravel. Yet they furnish the background for the development not only of modern French music but of modern European music in general.

LETA S. HOLLINGSWORTH. *The Psychology of Special Disability in Spelling.* Columbia University Contributions to Education, No. 88. New York: Teachers College, 1918. Pp. vi, 105. $1.40.

This is a kind of study of which the science of education has been greatly in need. The purpose was to observe for a considerable time pupils of normal general capacity but deficient in one particular school subject (spelling), to diagnose their disabilities by means of psychological tests, and to devise remedies for their defects. The pupils were selected from the fifth grade on the basis of reports by teachers, and tests with the Ayres spelling scale. The degree of intelligence was determined by the Stanford scale, an elaborate analysis was made of the exact nature of the spelling defects, and the effect of specific training was carefully noted. There is a valuable chapter on the theory of spelling defects, and a survey of the literature on the subject.

BENJAMIN KIDD. *The Science of Power.* With an Introduction by Franklin H. Giddings. New York: G. P. Putnam's Sons, 1918. Pp. viii, 318. $1.50.

This posthumous book by the well-known author of *Social Evolution* will arouse keen interest in thoughtful readers, and will stimulate animated discussion. On the basis of the early events of the war, and the writings of German and English leaders the effort is made to show the "failure of western knowledge," and the breakdown of ethical standards constructed upon it. In part two the author develops the thesis that the basis of power in civilization rests not on reason but on emotion. The passion for the ideal is the master fact of social integration. The key to this new "science of power," the psychic center to the new social integration is to be found in the nature and influence of woman. This daring and original solution is buttressed by skillful argument and copious reference to contemporary discussions. The growing political and social supremacy of woman in the western world lends itself to the support of this view. The book will be eagerly read by those interested in any form of feminism

T. SHARPER KNOWLSON. *Originality: A Popular Study of the Creative Mind.* Philadelphia: J. B. Lippincott Company, 1918. Pp. xvi, 304. $3.50.

This is a rather discursive and chatty consideration of the higher thought processes involved in creative activity. The author has read widely and makes thrifty use of the results in extensive quotations. In his own words his object is "first, to show the importance of a study of creative thought and to develop an interest in it; next to offer some suggestions as to the natural history of mind in its most inspired moments; finally, to institute a regime for the individual whereby he may secure the highest mental efficiency." Part I deals with the natural history of genius, and presents psychological, philosophical, and popular descriptions of the nature of consciousness, discusses Ostwald's view that consciousness is a form of energy, and gives some accounts of the relation between genius and the subconscious. Part II is concerned with the origin of new ideas, presenting various views of inspiration, deriving so-called laws of inspiration, and narrating instances of genius in pathological situations. Part III considers such biological factors as the relationship of age and sex to originality. Part IV enumerates hindrances to originality, as an undue regard for the past, defective home training, false education, lack of a science of reading, low standards of merit, incomplete effort, and professional bias. Part V examines the claims made by some that originality is now impossible, that the great steps in innovation have been taken, and that there is nothing left but minute amplifications of existing attainments. With this view the author has no sympathy, but rather sees indefinite expansion for the future. Finally, Part VI displays an array of reflections on mental attitudes and methods of creative activity. The particular merit of the book lies in the large number of opinions of eminent men brought together and focussed on the question of originality

ALBERT H. LEAKE. *The Vocational Education of Girls and Women.* New York: The Macmillan Company, 1918. Pp. xix, 430. $1.60.

In view of the rate at which women are coming to take the place of men in doing the work of the world it would seem to require no little daring to forecast just what vocational education girls should have. The author fortifies himself against any such attack by disclaiming in the preface any intention to consider the vast number of new occupations which women are now entering, and contents himself with de-

scribing the types of vocational education already offered. This task is admirably performed, and the book will undoubtedly take its place as a standard reference work on the subject. There is first a consideration of the status of instruction in household arts, and a description of typical courses for the elementary school, the high school, credit for work done in the home, forms of continuation schools, and special modifications for prevocational schools, homemaking schools and trade schools. There is a frank and fair discussion of the problem of domestic service and its possible solution in cooperative housekeeping. Part II deals with education for industries outside the home, taking up the conditions of such industries, the problems of factory labor, the organization of trade schools, evening schools, commercial education for salesmanship, and the possibilities of vocational guidance for women. There is an annotated bibliography of fifteen pages.

Maryland Public Schools. Fifty-first Annual Report of the State Board of Education Showing the Condition of the Public Schools of Maryland for the Year Ending July 31, 1917. Pp. 381.

This is an account of the operation of Maryland schools for the first year under the new school law. Features of the report are the workings of the compulsory school law, the methods of rural school supervision, and the status of rural school houses.

FLORENCE MATEER. *Child Behavior. A Critical and Experimental Study of Young Children by the Method of Conditioned Reflexes.* Boston: Richard G. Badger, 1918. Pp. 239. $2.00.

This is an interesting departure in child study. The author first gives a historical account of child study and a critical evaluation of its results. There follows a discussion of behaviorism and its significance for child study, and a rather full account of the experiments of Krasnogorski on the formation and breaking of conditioned reflexes in children. In her own experiments the author used a modified form of the Krasnogorski method, and supplemented this with physical measurements of height, weight, grip, lung capacity, and with tests of mental age by the Binet and the Yerkes scales, form board, and other tests. The author believes that the study of conditioned reflexes promises to yield important data for a real science of child study.

Measurement of Educational Products. The Seventh Yearbook of the National Society for the Study of Education, Part II. Prepared under the supervision of the National Association of the Directors of Educational Research. Bloomington, Illinois: Public School Publishing Company, 1918. Pp. 194. $0.90.

This yearbook constitutes a landmark in the history of educational measurements. It is full of interesting and valuable material that should not be ignored by progressive educators. The papers are "History and Present Status of Educational Measurements" by Leonard P. Ayres; "The Nature, Purposes and General Methods of Measurements of Educational Products" by Edward L. Thorndike; "Specific Uses of Measurement in the Solution of School Problems" by M. E. Haggerty; "General Organization of Measurement Work in City School Systems" by Frank W. Ballou; "Bureaus of Research in City School Systems" by Eugene A. Nifenecker; "Cooperative Work from a University Center" by Ernest J. Ashbaugh; "Existing Tests and Standards" by Walter S. Monroe; "Related Forms of Educational

Investigation" by W. A. Averill; "Statistical Terms and Methods" by B. R. Buckingham; "Training Courses in Educational Measurement" by S. A. Courtis; "Suggestions for Experimental Work" by George Melcher; "A Look Forward" by Charles H. Judd; and "A Selected Bibliography of Certain Phases of Educational Measurement" (606 titles) by Edna Bryner.

MILLICENT OLMSTED. *The Land of Never Was.* New York: Hurst and Company. Pp. 148.

This delightful story for the little folks describes a journey to Moon-Goose land, and recounts the experiences of the three children who took the trip. The essential feature of the visit was a school in which the pupils acted out all of the stories in the Mother Goose rhymes. It is excellent supplementary reading for the second grade.

MAURICE PARMELEE. *Criminology.* New York: The Macmillan Company, 1918. Pp. xiii, 522. $2.00.

In this book the well-known author of "The Science of Human Behavior" has attempted a comprehensive survey of the whole field of criminology. The book is a companion volume to his "Poverty and Social Progress," and the two together furnish a description of what he calls the two greatest social evils, poverty and crime. In tracing the origins of crime the author goes back to the equivalents of crime and punishment among animals, sketches the main types of crime among primitive peoples, and finds the key to its evolution in the conflict between the interests of the individual and organized society. Among the environmental factors involved in crime are climate and the weather, urban and rural populations, economic conditions, and political organization. On the subjective side we note the anatomical and physiological studies of criminals, the mental traits of the criminal diathesis, criminal amentia, various classifications of criminals into types, juvenile criminality and female criminality. Under criminal jurisprudence the author traces the development of criminal law, compares different types of criminal procedure, outlines the recent growth of recognized principles of evidence, advocates public defense as well as public prosecution, and indicates the functions of the judge, the jury and the police. Of particular interest for the psychologist is the discussion of the origin and evolution of punishment, the connection between moral ideas and punishment, the individualization of punishment, the arguments for and against capital punishment, and the evils of our prison system. Finally the author considers political and evolutive crime, freedom of speech, regulative social legislation, and the prevention of crime. Due account is taken of recent investigations in the psychology of crime, and there is a strong plea for a more general recognition of the psychological point of view.

AGNES LOW ROGERS. *Experimental Tests of Mathematical Ability and Their Prognostic Value.* Teachers College Contributions to Education No. 89. New York: Teachers College, Columbia University, 1918. Pp. 118.

This is an extremely important study of abilities in high school mathematics. There is a good summary of previous work on mathematical abilities, a description of the thirteen algebraic, geometrical and arithmetical tests used, and an account of the supplementary tests in language ability. The latter were the mixed relations, the logical opposites, the Trabue language scales, and the Thorndike reading scales.

The tests were given to fifty-three pupils of Wadleigh High School, New York City, and sixty-one pupils in the Horace Mann High School for Girls. The author presents elaborate tables of reliability coefficients and coefficients of correlation. It is interesting to note that correlation between verbal ability and mathematical ability is higher (.65) than that between any other pair of results. Whether this means that mathematical ability depends upon verbal ability or not, it indicates that the two abilities are found largely together. The diagnostic and prognostic values of certain tests were high, so that an hour and a half spent with these tests may be expected to give a correlation of .60 to .80 with future mathematical achievement.

HAROLD ORDWAY RUGG AND JOHN ROSCOE CLARK. *Scientific Method in the Reconstruction of Ninth-Grade Mathematics.* Supplementary Educational Monographs, Vol. II, No. 1, April, 1918. Chicago: University of Chicago Press, Pp. v, 189. $1.00.

This is one of the most important constructive monographs on the teaching of mathematics that has ever been written. In the introductory chapter the authors show the prevailing discontent with ninth-grade mathematics, and assert that it cannot be rearranged or reorganized but must be rebuilt from the ground up. There follow a minute analysis of what is done in ninth-grade mathematics, a historical study of how algebra became the ninth-grade course, an account of the construction of standard tests in algebra, a brief report on the application of these tests to many thousand pupils, an analysis of the errors revealed by these tests and the necessary corrective procedure, a study of the training in "logical thinking" afforded by mathematics, a survey of curricula in secondary mathematics, and an account of experimental teaching in ninth grade mathematics. The authors urge discussion, constructive criticism, experimentation, observation of others' work, the keeping of records, and co-operation amongst teachers.

MAJOR E. Z. STEEVER, III. AND MAJOR J. L. FRINK. *The Cadet Manual. Official Handbook for High School Volunteers of the United States.* Philadelphia: J. B. Lippincott Company, 1918. Pp. xxxi, 315. $1.50.

The High School Volunteers of the United States is the high school military organization officially recognized by the War Department, and it is now rapidly spreading over the entire country. The different parts of the Manual are devoted to the regulations, drill (taken from the infantry drill regulations of the United States Army), the use of the rifle, signals and signaling, different types of setting-up exercises, and wall scaling, cadet shows and competitions. An appendix gives the War Department requirements for installing the H. S. V. U. S. system in the high school.

C. W. SUTTON. *The Arithmetical Abilities of School Children as Shown by Courtis Tests.* Cleveland, Ohio: Division of Reference and Research, Bulletin No. 1, 1917. Pp. 15.

The results of the application of the Courtis Tests, Series B, to grades four to eight in thirty-six schools. Cleveland stands well in grades four, five and six, but falls off in grades seven and eight. Special drill is recommended for those children below standard. The danger here is in directing the attention of the teacher exclusively to the dullards and ignoring the more educable pupils. It perpetuates the vicious mass fallacy which educational measurements should help us destroy.

Third Report of the Committee on Economy of Time in Education. The Seventh Yearbook of the National Society for the Study of Education, Part I. Bloomington, Illinois; Public School Publishing Company, 1918. Pp. 134. $0.75.

The greater part of this year book is devoted to the place of history and civics in the elementary school. Other papers deal with social demands on the course of study in arithmetic, minimal essentials of elementary school geography, vocabularies of second year readers, and composition standards for the elementary grades. A history symposium emphasizes a broader outlook on world happenings and more attention to social and industrial conditions.

W. W. THEISEN. *An Educational Survey of Janesville, Wisconsin.* Madison, Wisconsin: State Department of Public Instruction, 1918. Pp. 329.

Of especial interest is Chapter XIII, *Measuring Results in School Subjects.* In arithmetic the Woody tests, the Courtis tests, Series B, and the Stone Reasoning tests were used. The Kansas Silent Reading Tests and the Ayres and Buckingham spelling lists were given, the hand writing measured by the Thorndike, Ayres and Freeman scales, compositions were evaluated on the basis of the Hillegas scale, and the Trabue Language Tests were given. Not the least important aspects of the survey are the valuable comments on methods of measurement.

ARTHUR JAMES TODD. *Theories of Social Progress. A Critical Study of the Attempts to Formulate the Conditions of Human Advance.* New York: The Macmillan Company, 1918. Pp. xii, 579. $2.25.

Is society progressing? In view of the aspect of the present orgy of destruction many will be inclined to deny it. Everything depends on what we mean by progress. The author contends that it should be possible to work out a series of objective tests for social progress; and if we are ever to have a Science of Society such objective tests must be part of its ultimate purpose. Thus the author stands committed to the same quest for objective measurements of social progress that some of us are engaged in for education. More than this,—he clearly recognizes how important for society as a whole is the striving for better and more efficient standards of education. "If humanity is to hold the threads of its own destiny and rise from ages of blind drift to a plane of mastery, it will be through discovering and utilizing new types of education." Frankly admitting that as yet we have no objective scale of social measurements, he thinks that the best method of attacking the problem is by surveying and organizing the theories of social progress that have been proposed. This he does in a very able and attractive fashion. documenting his argument with copious citations to the original sources. It is a sane, scholarly, and enlightening book, and deserves to be widely read by thoughtful educators.

WILLISTON WALKER. *A History of the Christian Church.* New York: Charles Scribner's Sons, 1918. Pp. xiii, 624. $3.00.

The broad-minded educator needs to be familiar with the history of human institutions of all sorts. No single institution has had such a profound educational influence, or embodies to so great an extent the highest ideals and aspirations of humanity as the Christian church. Perhaps we may not subscribe to many of the doctrines promulgated by the church in its various branches, indeed the present volume affords ample reason for the scientifically minded man to hestitate and ponder over religious dogmas of any sort, but every educated man should know the main

outlines of church history in order to appreciate the religious institutions of the present day and to judge of the modifications needed to meet the demands of the future. For this purpose the present book is admirable. The style is clear and graceful, as free as possible from technicalities, the author's scholarship is broad and well-grounded, and his presentation is frank, open, and free from bias. The book reads like a romance, and gives a thrilling picture of the development of western civilization. The value of the work is enhanced by many double-page maps.

A. N. Whitehead. *The Organization of Thought, Educational and Scientific.* Philadelphia: J. B. Lippincott Company, 1917. Pp. vii, 228.

These papers by one of the most eminent English mathematicians and scientists command the attention of all thoughtful students of education. Most of them are addresses delivered since the outbreak of the war, and as such reflect the sentiment of advanced English scientific thinking on the type of education demanded by the new conditions. While the author does not abate one tittle of his enthusiasm for and high valuation of mathematics, he recognizes frankly that mathematical training is for the few. And even for these few he urges a tremendous simplification and condensation on the vital aspects of the subject. Pupils are now overloaded with such a multiplicity of details that mathematics becomes isolated, detached, with little significance for the big things of life.

Robert Sessions Woodworth. *Dynamic Psychology.* New York: Columbia University Press, 1918. Pp. 210. $1.50.

This vigorous and forceful discussion of some of the fundamental problems of psychology will be of interest alike to the pure and to the applied psychologists. The author presents a historical sketch of the modern movement in psychology, distinguishes certain problems such as the difference between "drive" and "mechanism," examines the drives and mechanisms that are native to the human organism and those that are acquired, discusses the factors of selection and control in reactions and the factors of originality, and considers some outstanding drives and mechanisms involved in abnormal behavior and in the reactions of animals. By mechanism the author means the psycho-neuro-muscular adjustment on which a reaction depends, and gives an answer to the question how does the reaction take place. Drive, on the other hand, is the motive, the urge, the reason why the organism reacts as it does. Whether drive is not ultimately explainable in terms of mechanism is a question which the author is inclined to answer in the affirmative, but he believes that at this stage of our knowledge the distinction has a pragmatic value.

VOLUME IX SEPTEMBER, 1918 NUMBER 7

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

THE CLASSIFICATION OF PUPILS IN ELEMENTARY ALGEBRA

JAMES F. TAYLOR

Prin. Denfeld High School, Duluth, Minnesota.

INTRODUCTION

In all that pertains to intellect it is constantly illustrated that individuals differ in the ability to acquire knowledge. This truth is axiomatic. All know that some learn slowly, others rapidly. By consulting any good treatise on psychology one may find a useful discussion of these differences in mental traits and their causes; but one would not be so well rewarded if he should seek an account of the effects which these differences in ability to learn have individually and collectively on the members of a group.

In this article the latter phase of the matter will receive emphasis. I shall be concerned largely with such questions as:—(a) Is there any advantage in learning rapidly? (b) What effect does the slow learner have on the class work? (c) How much benefit does the slow pupil receive from the class work? (d) How much does the bright pupil receive? (e) What effects do the slow pupils have on the work of the bright pupils? (f) What effect does the bright pupil have on the slow one? (g) What are some of the conditions under which the best work for an individual can be accomplished? (h) What are some of the methods set forth for the solution of these problems?

My findings in some respects agree with the findings and opinions of others, but in several cases they are the reverse of popular and traditional opinion.

THE MATERIAL AND THE METHOD OF SECURING IT

In a large high school in Minnesota during the school year 1911–1912, the material for this article was gathered by the writer from his classes in beginning algebra as follows:

Each pupil was given a card upon which his or her daily record was kept according to the following items: (1) Name. (2) Date. (3) Period of recitation. (4) The daily assignment. (5) The number of problems solved correctly. (6) The number of problems solved incorrectly. (7) The time to the nearest five minute period spent outside of the recitation in preparing the lesson. (8) The grade in per cent. made on the assignment.

Several precautions were taken to secure normal and trustworthy data. The more important of these were:

(1) The pupils at no time understood that the material was to be used for the present purpose; but on the other hand they treated it as a part of their regular school work from which they would not profit except by getting the habit of checking their work, and learning how to use time to the best advantage in preparing lessons.

(2) The assignments were selected from the exercises in *Wentworth's New School Algebra*, and are accordingly numbered in this article. In the main these were drill exercises; they were assigned after the theory by which they were to be solved had been carefully studied in class. The theory was never assigned for an advance lesson. These precautions insured that each pupil knew how to proceed with the assignments.

(3) Each pupil was urged to work alone at a regular time, at a maximum rate of performance, and under his or her normal conditions.

(4) The time was kept to the nearest five minute period by the pupils. Uniformity in keeping correct time was insisted upon on the ground that they would learn to be more careful and accurate with their own work. Their honor was appealed to in this matter. The writer feels that the time element is as accurate and reliable as it is possible to obtain from normal and conscientious pupils.

(5) The pupils brought daily to the class written solutions of the problems. After these were graded by the class the daily record cards were made out and taken up along with the written work, which was afterwards graded by the writer to see that no mistakes had been made, and to secure uniformity in grading.

The Determining Factors in Selecting the Cases

The cases considered in this article are a selection of seventy out of one-hundred seventy pupils ranging in age from thirteen to nineteen years, representing many nationalities, and various economic and social conditions of home life. In fact the selection truly represents a cosmopolitan group of American high school pupils assembled from various wards and types of schools.

The agencies by which the selections were made were rather negative; some of them were: (a) Absence from class for a long period. (b) Failure to be regular in keeping records. (c) Failure to put forth a constant, and individual effort according to the pupil's ability.

Any one of these objections was sufficient to eliminate a case from the selection.

Special effort was taken to secure characteristic representatives of the work done in high schools. When a pupil missed a regular lesson he was given an extra lesson involving the same theory and required to make it up under the class restrictions. Thereby a very good record was obtained in the cases of absentees, and as the chronic absentees were eliminated from the selection, the few cases of absences included in this article would have no appreciable effect on the results of this paper.

The Treatment of the Material

In all there were sixty-six (66) exercises selected which are herein numbered accordingly. For each of these exercises each pupil handed in a daily record card from which a tabulation according to sequence of date was made showing the number of minutes spent and per cent. made on each assignment by each pupil. All the following tables and graphs are derived from this tabulation which for convenience we shall call reference tabulation, and which is too large to print. Great care was taken to make this table accurate.

Accuracy and Rapidity

Table I.

What relation is there between the very desirable traits of accuracy and rapidity? Nowadays we hear a great deal of discussion about them, most of which is not concrete.

Below is a table derived from the reference tabulation showing the classification as to grades, corresponding intervals of time, and distribution of numbers in each group.

	A or 100%	B or 86%–99%	C or below 86%
Total No. of cases making...	2234	1037	1349
Or reduced to per cent. of whole no. of grades.....	48.3%	22.4%	29.3%
Av. per cent...............	100%	92.1%	75.5%
Av. T....................	38.5 min.	47.1 min.	52.77 min.
Av. T′....................	38.5 "	51.2 "	77.98 "
Av. additional time of finishing the assignments.....	6 "	4.1 "	25.22 "

Av. T stands for the average actual time spent in preparation of the lessons outside the class by the pupils.

Av. T′ stands for the average actual time required for complete mastery of the assignments, supposing that the pupils continued at the rate determined by the work they actually accomplished.

A comparison of the various groups as to their accomplishments and time intervals will clearly show that the A group accomplished more in less time than either of the other groups, which definitely indicates that speed and accuracy go hand in hand.

In some of the exercises (8 in all) apparent variations are found. These variations can be easily accounted for on the ground that the poorer and slower pupils get tired and become discouraged in a shorter time than the more accurate and more rapid pupils when they have a long lesson whether easy or difficult. The poorer and slower pupils due to the lack of incentive and satisfaction accompanying the overcoming of difficulties, become satisfied to stop work after they have spent a certain amount of time, or have obtained solutions for a certain portion of their assignments. In fact they become satisfied with a lack of mastery, and set for their goal the passing mark, which is a very damaging attitude for mental growth. Guided by this motive and standard the poorer pupils consciously or unconsciously set a maximum time over which they will not work. From this we readily see that these apparent variation are not real exceptions, but they too emphasize the same thing that the other exercises emphasize.

Instead of comparing the T columns we should compare the T′ columns; at once it would be apparent that the C group did not show any ability to master the assignment in a shorter time than either of the other groups. The A and B groups set for their goal the mastery of the lesson. Thus they will persevere until they have either mastered or completely solved the problems assigned, regardless of the length of time required. From this it is readily

seen why the brighter pupils will, if necessary, put in more time to master the work than the poorer will in order to partially master it. The point might be made clearer by an illustration. For instance take a difficult exercise as 67, in which the A and B groups each on an average spent 76 minutes, while the C group spent 70 minutes. This is a significant fact when it is noticed that the C group were satisfied with their efforts at the end of 70 minutes, whereas it would have taken them at their rate of work 93 minutes to have mastered the assignment. They were ready to stop 23 minutes before they should have, while those in the A and B groups in order to completely master the work were willing to continue 6

TABLE II

	A.		B.				C.			
Time in 5 min. intervals.	Distribution of cases according to actual time spent in prep.		Distribution of cases according to actual time spent in prep.		Distribution of cases according to the time required to completely master the assignments		Distribution of cases according to actual time spent in prep.		Distribution of cases according to the time required to completely master the assignments	
	(1) cases	(2) av. time	(3) cases	(4) av. time	(5) cases	(6) av. time	(7) cases	(8) av. time	(9) cases	(10) av. time
20–25	49	22	0	0	0	0	0	0	0	0
26–30	47	27	14	29	14	29	12	28	0	0
31–35	40	33	15	34	11	32	11	33	12	39
36–40	30	38	15	37	17	38	11	38	10	43
41–45	29	43	14	42	15	43	17	43	9	47
46–50	27	47	19	47	16	48	23	49	12	53
51–55	26	53	18	54	18	53	22	52	16	59
56–60	20	58	13	58	19	59	25	58	18	64
61–65	21	63	16	63	12	63	28	64	21	69
66–70	17	66	0	0	16	68	28	68	18	73
71–75	24	74	18	73	18	72	20	71	32	79
76–80	24	76	14	76	18	79	0	0	26	83
81–85	0	0	6	83	14	84	0	0	28	89
86–90	0	0	0	0	0	0	0	0	25	92
91–95	0	0	0	0	0	0	0	0	29	101
96–10	0	0	0	0	0	0	0	0	0	0
100–105	0	0	0	0	0	0	0	0	0	0

minutes longer than the C group. Evidently we are safe in concluding that the most accurate pupils are the most rapid performers, and also their maximum length of time for accurate work is greater than that of the slower and less accurate.

It should be noticed as significant that pupils are capable of doing accurate work at their maximum rate for a period of 75 or 76 minutes. It will not be understood that a beginner could work for this length of time from the first, but as he advances in a study the length of time at a maximum rate of work can be increased by practice.

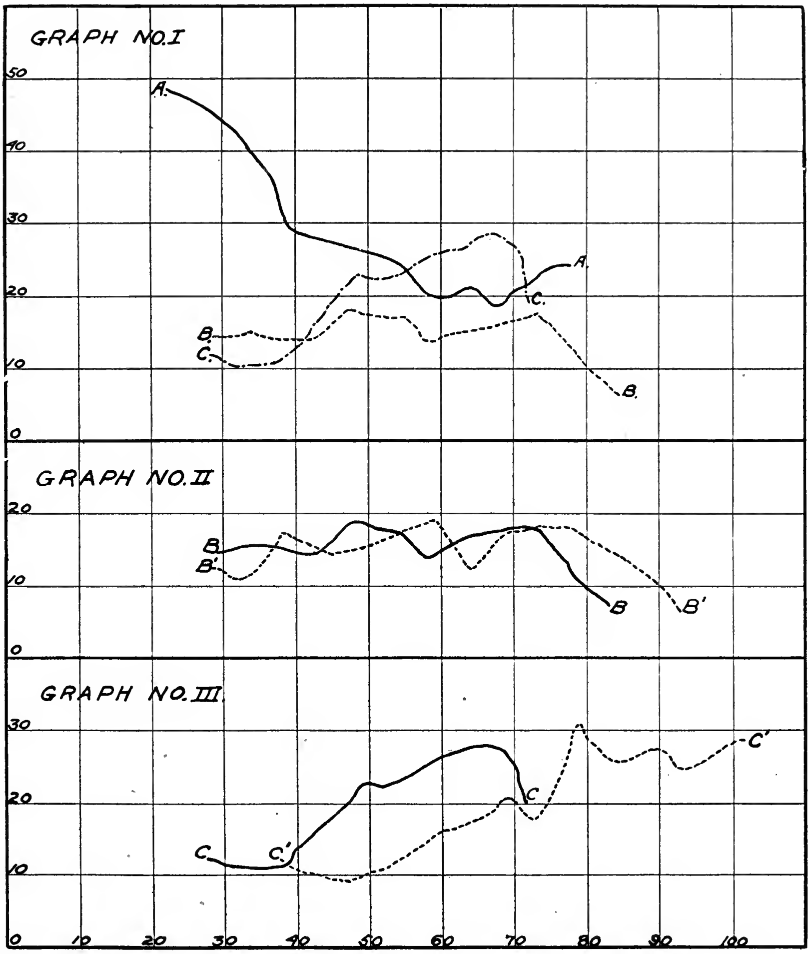
THE EXPLANATION OF THE FORMATION OF TABLE II

Table II is a tabulation of the average number of cases in each of the groups at an average time taken in intervals of five minutes each. In order to clarify, an illustration is here inserted. From the reference tabulation all the cases in the A group who were able to do their assignments within the interval between 20 and 25 minutes are found in:

(1) Ex. 9—54 cases with an average time of 20 minutes.
(2) " 15—42 " " " " " " 25 "
(3) " 39—45 " " " " " " 22 "
(4) " 42—54 " " " " " " 22 "

Thus for the four exercises above considered there are 105 cases, or an average of 49 cases for each exercise, and for the same exercises a total time of 89 minutes, or an average of 22 minutes for each. For the three groups the same classification was carried out and respectively tabulated under the proper headings as illustrated above.

Take for instance the time interval 36—40 minutes in the A group, Table II; we find 30 cases in column marked (1) with an average time of 38 minutes in column (2). Since the A group completely mastered the assignments in the preparation, the columns indicating required time for a complete mastery are the same as those for preparation. For the same interval there are under the B group 15 cases in column marked (3) with an average time of 37 minutes for their actual preparation in column marked (4), while for the required time for a complete mastery there should have been for this interval 17 cases in column marked (5) with an average of 38 minutes in column marked (6). In the C group there are 11 cases in column marked (7) with an average time of 38 minutes for actual preparation in column marked (8), while there should



In Graph 1 the distribution of groups A, B and C are indicated according to their actual performance. In Graph 2, B indicates the distribution of group B according to actual performance and B′ the same group according to required performance for complete mastery. In Graph 3, C indicates actual performance and C′ the required performance for complete mastery.

have been for this interval 12 cases in column marked (9) with an average time of 39 minutes for complete mastery in column marked (10).

The Variations in Time and Per Cent. as Shown by Graphs 1, 2, and 3

Graph 1 represents the distribution of the cases for the A, B, and C groups respectively according to the actual length of time spent on the assignments. For each branch time is used as the abscissas, and the number of cases as ordinates as they are tabulated in Table II in the Columns (2), (4) and (8), (1), (3), and (7). By taking the extreme abscissas of the A and C branches and comparing the differences it is clearly evident that the range of time of work for the A group is greater than that for the C group. This indicates that the C group gives up sooner than the A group, which has already been pointed out above. With this understanding we see what a difficult and important task it is to assign a lesson. A maximum assignment for the A group could not in any manner be mastered by the C group. The assignments must not be more than the maximum amount the poorer pupils are willing or disposed to do, provided the assignments are to be completely mastered by them, which in general is one of the most important factors in determining the length of assignments.

From the general shape of the A branch of Graph I (that is, that the abscissas do not decrease as rapidly as the ordinates increase) it is clearly indicated that the time of accurate work can be increased by practice. Take for instance in Ex. 20 where twenty-one made A grades with an average time of forty-six minutes, while in Ex. 92 after several months study thirty-three made A grades with an average time of forty-five minutes.

Graphs II and III are made up of the B and C branches of Graph I, together with branches made in a similar manner by using columns (5), (6), (9), and (10) of Table II for the added branches, columns (5) and (9) being used for ordinates and (6) and (10) for abscissas.

It must be borne in mind that there is not a one-to-one relation between the branches of these curves; for it has been shown that on an average the B group should have worked about five minutes, and the C group about twenty-five minutes longer in order to have completely mastered the work assigned. That is, in Graph III

an abscissa of 20 in the branch of actual performance would correspond to an abscissa of 45 in the branch of required performance.

After making due allowance for this variation the general shape of the curves in Graph III clearly indicates that the ordinates do not change much as the abscissas vary from thirty minutes up to forty minutes, a nearly uniform variation in the ordinate from forty minutes up to seventy minutes, and from there on the ordinates remain almost constant as the abscissas increase.

This in turn means that those pupils who are slow and inaccurate after a time become more and more fixed in the C group, and as the lessons or assignments increase in length or difficulty the less accurate they are until finally they become permanently fixed in the C group.

After making due allowance for the variations in the abscissas in the branches of Graph II, it is manifest from the slight variation of the ordinates as the abscissas take on various values from the smallest to the largest, that there is not much of a change in the B group except a slight shift from the A to the B, and just about an equal shift from the B to the C group according to the difficulty of the assignments.

Graph III also shows, when compared with I and II that the poorer become poorer more rapidly than the brighter improve, because the range of required time to master the assignments as time passes becomes longer and longer more rapidly for the C group than it does for the other groups.

The knowledge of these truths is not new to current lore, but their true significance has not attracted much serious attention for improvement in our actual school practice.

Everybody knows that when a team is overloaded the more the wagon is agitated by jerks, the deeper it sinks in the mire. Understanding this the teamster at once lightens the load and takes lighter loads. This could readily be done in school work provided all were in the mire or there were only one or two pupils to deal with. But when there are about twenty-five or thirty pupils with about as many different abilities it becomes a very different and difficult problem. It becomes a serious matter, when we view the school work from this angle, that the brightest pupils have no opportunity to work on assignments comparable to their abilities and that the poorer pupils set the standards and will not do any more than they are willing or disposed to do. The question of adjusting

the school work so that each pupil may receive the most good, becomes quite serious when we stop to ask how much should the brighter and by far the larger proportion of the pupils be hampered for the slower and less capable who profit little or none even by the assignment of work commensurable to their abilities.

The Effect of Grouping the poor and Bright Pupils in the Same Class

It has often been said that "Poor or dull pupils become animated and aroused in their studies by being associated with brighter pupils in classes." This assertion implies a beneficial relation existing between these two groups of pupils, or at least enough benefit is received by the poorer to over-balance the loss to the brighter by such classifications. Let us determine whether or not such a statement is well founded. Let us take up the question of effects the A, B and C group have on each other, when they are placed together in the same classes.

TABLE III

Groups	Av. Time Worked	Av. Time Required to make 100%	Per Cent made	Additional Time required to make 100%
A, or 48.3% of pupils	38.5 min.	38.5	100	0
B, or 22.4% of pupils	47.1 "	51.2	92.9	4.1 "
C, or 29.3% of pupils	52.77 "	77.98	75.5	25.21 "

From the above table it is evident that the A pupils do not require any additional time in order to correct mistakes made in their preparation. Therefore they do not necessarily receive any benefit from that part of the recitation designed for correcting errors. In the case of the B group since they did not completely master the assignments in their preparations additional time is required by them for correcting errors. It is noticed that this time is 4.1 minutes. Thus on the average the B group of pupils receive benefit for at least four and one-tenth (4.1) minutes from that portion of the class work designed for correcting errors and mistakes. In the case of the C group it is also noticed they failed to master their assignments and that an average of 25.21 minutes is required by them to complete the assignments and to correct errors made in preparation.

That is the time required, supposing the class work can be carried on at the same rate as that determined by the C group's preparation, which rate, by the way, is the slowest of the three groups. This class rate is not very likely to be equal to the preparation rate, because the wrong notions that have been formed must not only be eradicated, but new and correct ones must be formed to take their places. This is a much harder and slower process than to form the original conceptions correctly.

In so far as our American schools are organized on the assumption that each child shall be granted the right to have all his questions answered as well as time and a chance to completely master his work, it is clear how much the C group has a right to ask and expect.

We see then that the A group, 48.3 per cent. of the pupils, the brightest and most capable, has at least twenty-five minutes of its recitation time wasted, which is five eighths or 62½ per cent. of a forty minute recitation period. The B group, or 22.4 per cent. of the pupils, fares a little better. As the B group fails to completely master the assignment at least 4.1 minutes or about 10 per cent. of the recitation should be set apart for correcting errors. After this is done there is at least 20 out of 25 minutes set apart for corrections or 50 per cent. of the recitation period, from which the B group receives no benefit. The C group, alone, the poorer, the slower and the less capable, 29.3 per cent. of the pupils, has a chance to receive benefit for the full 25.21 minutes, 62½ per cent. of the forty minute recitation period, which the C group alone demands should be set apart for that purpose.

Thereby the only group of pupils that has a chance to work to a maximum, and it has been shown above they will not, is the C group, which in general does not have the mental energy, courage, desire or enthusiasm to inspire them. Thus ere long the work becomes a drudgery, and as a result this group drops deeper and deeper into the intricacies of the work with no hope of extraction.

It takes no argument to show that the brighter and more capable pupils lose daily from 20 to 25 minutes of a recitation, and that instead of receiving benefit for the full 40 minutes they benefit for 20 to 15 minutes. Nor is the loss of recitation time the total time loss to them. It has been pointed out above that they are able to continue accurate work at their maximum rate for a period of at least 76 minutes, and it has also been pointed out that their interval of voluntary work is longer than that of the C group which cannot

be urged to continue work more than about 52 minutes, which time is, in rough terms, about $^2/_3$ of 76 minutes or 66 $^2/_3$ per cent. of the length of time the brighter pupils will voluntarily work in order to master their assignments. Therefore we conclude that the brighter pupils are also handicapped in the amount of work they could and would willingly do in assignments if given an opportunity.

To be sure no one would likely hold that an assignment should require, under ordinary High School conditions, an hour and a quarter for preparation, but it is not an extreme either way to hold that the assignments should be sufficient in length or difficulty to keep the brighter pupils busy at their macimum rate of performance for at least 50 or 52 minutes, which is the average length of time the C group will work. Thus it is evident there is a loss daily in preparation of about fourteen (14) minutes to the A pupils and about five (5) minutes to the B pupils on the basis that the A pupils work 38 minutes and the B group 47 minutes on the assignments. Counting the loss to the A group in recitation 25 minutes and on assignments 14 minutes per day, we find that the brightest and most capable pupils lose daily 39 minutes, while the B group loses 20 minutes in recitation and 5 minutes on assignments which is a total loss of 25 minutes. On the basis that a pupil including preparation and recitation, should spend daily at least 90 minutes work at his or her maximum rate of performance, we see that the A and B groups are so hampered that they are not able to put in more than about $^2/_3$ or 66 $^2/_3$ per cent. of that time. To be accurate the A group, 48.3 per cent. of the pupils, is held back daily to an extent of 39–90 of the time, almost 4–9, or 44.4 per cent. of the time allotted to a subject, and the B group 25–90, 5–18 or 27.7 per cent. of the same time. From the above we are now safe in concluding that more than 33 1–3 per cent. of the time of 70.7 per cent. of the pupils is sacrificed for the 29.3 per cent. of the pupils.

Now after the sacrifice has been made, what are the final results? To be sure it does not make the C group anything other than the C group, for it has been pointed out above that they become more fixed in that station the longer they continue in such class conditions. Therefore, it is apparent that the popular opinion "That a poor pupil will receive inspiration and encouragement by being put in a class with bright pupils" is not well founded. The general rule appears to be that they do not so receive an uplift; but rather that they do not attain as much as they are actually able to do, notwithstanding the loss to the A and B groups.

In General the Poorer Pupils do not Profit by Being Put in Classes with Brighter Pupils

The following Table is a tabulation for ten cases selected according to the lowest examination grades.

TABLE IV

No. of case	Av. per cent. made on preparation	T	T′	Final test grades
	1.	2.	3.	4.
18	92	45	49	78
21	87	43	48	70
27	93	51	55	76
34	83	70	84	75
43	73	42	58	75
46	80	44	55	73
48	74	45	61	69
49	78	38	48	75
64	84	41	49	73
65	79	45	56	75
General Averages...	82.3	46.4	56.3	73.9

In a general way this selection represents the C group, and whatever is deduced from it will be true of the C group as a whole. It will be noticed at once; (1) that the preparation grades average 82.3 per cent. after they had spent on an average 46.4 minutes; (2) that these pupils on an average should have spent about 56.3 minutes in preparing their assignments. That in turn means 10 minutes, the difference between 46.4 and 56.3 minutes, should have been set apart in class for correcting mistakes made in preparation. In so far as 25 minutes per recitation were spent in correcting errors, it is evident that more time was spent for that purpose than was demanded by this particular selection. Under such a condition it would be fair to expect that this selection should have profited by that part of the recitation, and that their examination grades should indicate the benefit derived by being greater than their preparation grades, but contrary to expectation, the average examination grade shows an average decrease of 8.4 per cent. (82.3 per cent.—73.9 per cent.) from the average preparation grade.

This result points to the fact that the poorer pupils do not profit by that part of the class work given over to correcting errors,

as much as is generally supposed. The reason that they do not profit is very probably that the poorer pupils try to do work that is not commensurate with their normal abilities. This abnormal strain, under which they work, causes them to make a large number of improper associations. In the recitation these improper associations are not eradicated as is usually supposed. These wrong associations which have not been thoroughly corrected keep coming up in the mind, when the true associations are presented, until the pupils become more and more confused. This confusion causes them to rely on bare memory of the facts to carry them through, and as bare memory is not reliable, failure to master the subject is invited as the final result.

The points of this discussion emphasize the statement that poor pupils, in general, do not gain anything by being grouped with the brighter and more capable pupils.

The Most Capable as Well as the Less Capable Pupils Losf By the Present Prevalent Method of Classification

By use of Table V it is sought to show how the brighter pupils are retarded in their school work when put in classes with poorer and slower pupils.

In the following discussion it is assumed:

(1) That a grade made in a preparation (called a preparation grade) is a fair measure of a pupil's normal ability to acquire a mastery of a subject when working alone.

(2) That the grade made on examination, after the subject had been thoroughly discussed, is a measure of a pupil's actual mastery.

(3) That the difference between the examination grade and the preparation grade indicates whether or not a pupil's actual mastery is equivalent to his normal ability to acquire.

Table V is a tabulation of the examination grades and the preparation grades for the four selections, each consisting of twenty-three (23) cases. The fourth selection marked (w) comprises the twenty-three cases who made the lowest preparation grades.

From the general averages of Table V it is readily seen that the preparation grade for the selection marked (x) is 94.4 per cent., and the examination grade for the same selection is 96.8 per cent. which is an increase of 2.4 per cent. over the preparation grade. Since the examination grade is the larger, it is evident that this

TABLE V

(x)			(y)			(z)			(w)		
Case No.	Exam. grade	Prep. grade	Case No.	Exam. grade	Prep. grade	Case No.	Exam. grade	Prep. grade	Case No.	Exam. grade	Prep. grade
2	100	98	47	90	99	48	69	74	30	90	71
20	100	95	2	100	98	21	70	87	43	75	73
58	100	98	3	93	98	64	73	84	48	69	74
22	98	95	15	95	98	46	73	80	40	80	75
25	98	96	26	92	98	34	75	83	44	83	75
28	98	91	58	100	98	43	75	73	49	75	78
41	98	94	69	98	98	49	75	78	65	75	79
52	98	95	18	97	97	65	75	79	23	92	80
57	98	92	33	85	97	27	76	93	46	73	80
69	98	98	4	97	96	13	78	92	29	90	81
4	97	96	14	92	96	50	79	84	36	90	81
18	97	97	24	95	96	16	80	88	37	92	82
8	96	91	25	98	96	95	80	84	39	83	82
63	96	87	53	91	96	40	80	75	34	75	83
15	95	98	59	91	96	42	80	92	35	80	84
24	95	96	1	90	95	67	80	89	50	79	84
31	95	92	5	83	95	5	83	95	64	73	84
61	95	92	20	100	95	6	83	91	62	83	85
11	94	93	22	98	95	19	83	90	38	90	86
3	93	98	52	98	95	39	83	82	21	70	87
56	93	90	41	98	94	44	83	75	63	96	87
14	92	96	10	83	93	54	83	89	16	80	88
23	92	80	11	94	93	62	83	85	45	88	88
Gen. Av. %	96.8	94.4	Gen. Av. %	94.3	97.9	Gen. Av. %	78.2	84.4	Gen. Av. %	81.8	81.1

selection evidently received benefit from the recitation and teacher. Therefore, this increase is the measure of the amount of good the (x) selection derived from the class discussions and the individual help given by the teacher.

The preparation grade for the (y) selection is 97.9 per cent., which is an increase of 3.5 per cent. over the preparation grade of the (x) selection. This difference indicates that the (y) selection has ability to more nearly master a subject alone than the (x) selection. Judging from their preparation records the (y) selection in ability is ranked first and the (x) selection second.

Again from the general averages, Table V, the preparation grade for the (y) selection is 97.9 per cent., and the examination grade for the same selection is 94.3 per cent., which is a decrease of 3.6 per cent. from the preparation grade. It is clearly evident that the (y) selection did not receive any benefit from the recitation. This difference of 3.6 per cent. as in the case of (x) selection, is the measure of the amount of good the (y) selection, which represents the pupils of first rank ability, received from the class discussions and individual help given by the teacher. It is then sufficiently manifest that the brighter and most capable pupils are not assisted but rather retarded in their studies, by that part of the recitation designed for correction of errors and misconceptions.

From general averages of Table V it will be noticed that the examination grades for the (w) and (z) selections are lower than the preparation grades. Therefore, the same conclusions in regard to class corrections are drawn for these selections as were drawn for the (y) selection.

Let us now compare the examination grades of the (y) and (x) selections. For the former the grade is 94.3 per cent. and for the latter 96.8 per cent. which is an increase of 2.5 per cent. for the (x) selection over the (y) selection. This difference indicates that the (x) selection more nearly mastered the subject than the (y) selection. Let us not forget the (y) selection, according to the preparation grade, has more ability than the (x) selection, but on the other hand, judging according to the examination grade, which is the measure of the amount of the subject actually mastered, it is found that the (x) selection shows that it, in truth, more nearly mastered the subject than the (y) selection, which represents the more capable pupils. Therefore, we can safely conclude that these pupils with the most innate ability and from whom we should expect the best

and highest results, judged by the amount they really master, which is indicate by their examination grades, fail to attain first rank, but instead take second rank with the first rate ability, while the B pupils with second rate ability attain first rank in the actual amount mastered.

A satisfactory answer for this condition can be found by looking back and calling to mind that the A pupils were hampered in the following ways:

(1) They were not given a chance or an opportunity to do work adapted to their abilities.

(2) In the recitation they are idle or inactive about 62½ per cent. of the time.

(3) They were given assignments that were at least 27 per cent. less than they should have been, if they had been given work enough to have required 52 minutes daily to prepare it.

Thus, we find the A pupils with much spare time, which coupled with the fact that they are never required to use their full ability, soon leads them to the consciousness that not more than a half effort is required of them in order to keep up their work. Under these influences they are led by distracting conditions to procrastinate, so that ere long a link is dropped, and then another. Although they think they will make amends in due time, but in time of need the amends are never made, and as a final result they are forced to take second place, while the B pupils, who are at all times conscious that they have a task that requires constant conscientious normal action to almost their full ability keep alert for everything that is necessary for a complete mastery, and they, ere long take first rank with second rate ability. Another contributing factor in favor of the B group is, that they more nearly have work that calls forth maximum ability than the A group.

In so far as the C group is granted the privilege of setting the standards for the work; (1) it is evident that not only the B group is hampered, but also the A group is so hindered and retarded by the arrangement that they are reduced to a second rank instead of a first rank standing; (2) it has been proved that even the C group does not receive all the good that they were given a chance to accomplish under such an arrangement; (3) the B pupils profited most from such a classification, and from the fact that they maintained certain wholesome attitudes toward their work, we may conclude that the most efficient mental growth at all times is attended by conscious continuous maximum of efforts and an alert-

ness for the facts and principles of the subject matter, which are necessary for its mastery. When such is not the case, as is with the A and C pupils, the best results cannot be attained.

With these truths in mind, it becomes little less than criminal to put pupils with various abilities together in classes, where 70 per cent., the A and B groups, are positively retarded in their mental growth by being hampered by the standards of the 30 per cent., the C group, who are not in any certain degree helped by such an artificial classification.

Summary, Suggestions and Comments

(1.) The most rapid performers are the most accurate.
(2.) The macimum length of effort of the most accurate and most rapid pupils is greater than of the slower and less accurate.
(3.) The slower and less accurate pupils become more and more fixed in the poorer group.
(4.) The length of time of accurate performance can be increased by practice.
(5.) The poorer group is the only one that has an opportunity to work to their full ability.
(6.) The brighter pupils are retarded for the sake of the poorer pupils.
(7.) The best retainers are the most rapid workers.
(8.) The best retainers are those who most nearly work with a maximum of effort at a maximum rate.
(9.) The best retainers are not necessarily the most thorough pupils.
(10.) It often happens that the most capable pupils take second rank, while first rank is attained by those of lower ability.
(11.) In general it does not hold that a poor pupil will be greatly benefited by being placed in classes with brighter pupils.
(12.) The most efficient mental growth must be accompanied with a conscious continuous maximum of effort, and alertness for facts and principles necessary for a mastery of the subject.

The problem of so arranging the school schedule that each pupil may be able to work to his full ability is not new. Every school system in some form or other consciously or unconsciously makes a more or less satisfactory attempt to solve it. It would be useless to try to enumerate all the methods and devices proposed for its

solution; however, let it be sufficient to call to mind the importance of the problem by noticing some of the more common, which are more or less make-shifts and not solutions.

(1.) By allowing the bright and more capable pupils to carry extra work.
(2.) By excusing those above a certain standard from examinations.
(3.) By organizing classes for repeaters.
(4.) By giving supplementary work.
(5.) By making the course so that the brighter pupils can finish in a shorter time.
(6.) By giving individual help to the pupils.
(7.) By furnishing tutors for the poorer and less capable.
(8.) By organizing the work along certain lines according to certain methods. This device has so recently been in vogue that all readily recognize it.
(9.) By writing text books so as to give a minimum and a maximum course.
(10.) By organizing the work so that it attracts the pupils.
(11.) By organizing classes for retarded pupils.

From these illustrations it is plainly evident that this problem is generally recognized in all phases of school life. No one as yet has been able to propose a satisfactory solution; however, some plans may be better than others. Among the best might be mentioned a system, where the work is so organized that a pupil, if capable, can do the work in a two term course, or if slow be allowed three terms.

In some places the course is organized to develop leaders from the brighter class who set the standards high. In order to keep the balance between the brighter and poorer students, tutors are furnished to assist the teacher with his work. Since the course is set for the brighter pupils, the poorer pupils must take care of themselves or be dropped. By this system something like two years of work is gained by the time the High School Course is completed.

Some schools endeavor to give individual instructions to each pupil. By this plan each pupil has an opportunity to do his maximum work, which is very good. No doubt each device proposed has one or several good points, but the satisfactory solution is not obtained. There may not be any such solution. But be that as

it may, there undoubtedly should be something done to stop the wail that is going up from all quarters, "What shall be done with the bright pupil?"

In the face of the above proved facts it appears that some relief might be had if the school should seriously recognize the difference of pupils' abilities and rates of performance, and organize the course along such a line. Let the false standard, that all pupils have the same ability, be broken down, and for once truly and frankly recognize that they are wholly different, and thereby cannot uniformly perform the same task at the same rate. It appears that until such a scheme can be wrought out, the school will go on sacrificing the best and by far the greatest amount of talent for the 30 per cent., the poorer, which sacrifice does not necessarily bring any good to them. Each pupil should be situated so that he could and would put forth a conscious continuous maximum of effort and alertness, in order that the best mental growth be obtained.

WHOLE VERSUS PART METHODS IN LEARNING NONSENSICAL SYLLABLES*

L. A. PECHSTEIN

University of Rochester

Whole method procedure demands the continuous repetition of an entire body of material until the desired stage of mastery is attained. Part procedure demands an initial mastery of definite sections of the material and the final connection of these different sections in proper serial order. It has seemed established that the whole method proves the more efficient in the field of learning verbatim. A recent paper† has shown, however, that certain modified forms of the part method prove far superior to the whole method when a motor problem is being learned. In so far as the maze problem may be considered typical of the motor field, one is led to face one of several propositions: (1) There is no correlation to be drawn between learning upon the motor and ideational levels. This proposition is substantiated by no experimental evidence. (2) The results obtained for motor learning are inapplicable to the ideational realm, since the motor problem studied allowed the subjects to distribute their learning over many days, whereas the accepted results for learning verbatim have in general been secured under massed conditions of learning effort. The validity of this proposition is sound until the relationship between distributed and massed conditions of learning is thoroughly established. (3) The field of learning verbatim is not sufficiently tested. This statement has validity only in so far as it suggests that certain modified forms of the part method have not been employed in this field and that final generalization as to the superiority of the whole method must be deferred until these modified part methods have had opportunity to show that they fail to duplicate the efficiency shown in motor learning. This paper directs attention, therefore, primarily to learning nonsensical syllables by numerous whole and part learning methods and secondarily to comparing the results with those for learning motor problems under like massed conditions.

---

*Read at the annual meeting of the American Psychological Association, December 27–29, 1917.

†PECHSTEIN, L. A., *Alleged Elements of Waste in Learning a Motor Problem by the Part Method*, Jr. Ed. Psychol., 1917, 8, 303.

The experimentation employed two series of nonsensical syllables. The A–Series consisted of thirty-two syllables in consecutive arrangement. The B–Series consisted of the same syllables but arranged in sixteen pairs, as 1–2, 3–4, 5–6, etc. The series were mounted on a mechanically driven drum and were exposed singly or in pairs through an adjustable aperture. The A–Series required an exposure of two seconds per syllable, the B–Series four seconds for each pair. The A–Series was considered learned when the subject could successfully anticipate and write each syllable when shown the immediately preceding one. The B–Series was considered learned when the subject could write the second syllable of each pair when the first syllables were shown as stimuli. These stimulus words were arranged in different order from their position in the original paired learning series. Each subject was given a learning trial and a test for measuring the amount learned immediately after each such trial, two seconds being allowed to anticipate each subsequent syllable in Series A or to recall the paired associate in Series B. The amount learned was calculated for each trial and the difference between a perfect score and this amount learned per trial was called the learning error.* Students in introductory psychology were used for subjects, no student having had earlier experience with the learning material. Each student learned one series by an assigned method. Six students constituted a learning group for a single method and series and the results obtained were averaged.

A description of the methods employed reveals significant differences. Of the whole methods, the learner may be allowed full opportunity to retrace the material earlier traversed in the trial and hereby to work over certain difficult sections before reaching the end of the series. This method is apparently a natural one for some organisms and is designated the Whole Method—'Returns Allowed,' Type A. The learning series is kept upon the drum, whose rotation is controlled as desired by the learner. The method renders it possible to keep a record of the number of times the material is completely worked through and to test the amount learned after each such trial. Because of the unlimited possibilities of returning, a question immediately arises regarding the satisfactory criterion of learning. The student soon learned not to complete

*A syllable was considered learned when it was spelled correctly and recalled in ts proper place. No credit was given if the syllable was out of place. A single letter was credited if properly placed.

a learning trial until he felt reasonably sure of having mastered much of the material. Consequently, the number of trials for learning and the number of errors would seem so small as to argue that the method is quite efficient. On the contrary, the excessive amount of time consumed by each trial argues for the inefficiency of the method. The time measurement seems the only just one for comparative purposes. Type B differs from Type A in that the student is allowed to completely learn the material in private before being tested, with retracing as desired. Time is the only measurement of the learning. Contrasting with these whole methods which allow full opportunities for returning, the method entitled 'Returns Prevented' requires the subject to study each syllable a constant period of time and to pass directly to the next one of the series.

The part methods are four in number. The "pure part" method requires that the material be divided into four equal parts, each part learned as a unit, and then the connecting of the units in proper serial order. For the A–Series, this makes each unit to consist of eight syllables; for the B–Series, of four pairs of syllables. The "progressive part" method requires that the first two sections be learned as separate units and then connected as soon as the second is mastered; the third unit is then learned separately and immediately added to the earlier material; the fourth unit, learned and then added to the complex unity already established, completes the learning act. The "direct repetitive" method provides that the first section be learned as a unit and that the learner review this section at each trial as an introduction to the new second division. When the first two units are finally thus mastered, they constitute the introductory review as section III is being learned, and the three in turn for section IV. The "reversed repetitive" method permits as the initital learning the last unit of the series, whereupon the subject is introduced to the new next-to-the last section, reviewing the previously learned one as the final part of each learning trial. When these two units are mastered, the subject is introduced to the next earlier unit and reviews all the previously learned sections as the final part of each learning trial. These two methods are, therefore, review methods and differ only as regards whether the review introduces or follows the unlearned unit. The "progressive part" and the two "repetitive" methods agree with the "pure part" method in satisfying the requirement that the material be

learned in parts; they differ from the "pure part" method in that they significantly modify the time and manner of connecting the several separately learned parts.*

These methods lend themselves most admirably for employment in learning a maze earlier reported.† This maze is divided into four distinct parts or units, each possessing the same number and type of cul de sacs, the same linear distance in the true pathway, entrances and exits into a common place, etc. By the removal of sliding panels, the maze can be learned as a whole or separately learned parts easily connected. The number of trials, the errors (*i. e.*, entering the blind alleys), and the total learning time can easily be measured. Consequently, learning this maze is highly comparable to learning such material as the nonsensical series A, when the sections are equated for length, and where the material may be learned as a whole, in parts or in any desired combination of parts. It is slightly less comparable to the B–series, for this series requires the learner to state the proper associates in an order differing from that of the original learning and hence renders impossible the fixing of such definite unit habits as are allowed and required by both the A–Series and the maze.

---

*The time and manner of the complex act of connection are exceedingly important. Positional relationships of space and time are established during the learning of each unit and these are markedly reconstructed during the acts of connection. For an analysis of these relationships and their significance in whole-part learning, see PECHSTEIN, L. A., *Whole vs. Part Method in Motor Learning. A Comparative Study*, Psychol. Rev. Mon. Supp., Vol. xxiii, No. 2, p. 42.

†Ibid. pp. 4–5; 70; Also, Jr. Ed. Psychol., 1917, 8, p. 304.

TABLE I

*Whole and Part Methods in Learning**

Method	Trials			Errors			Time		
	Maze	A–Series	B–Series	Maze	A–Series	B–Series	Maze	A–Series	B–Series
Whole									
Returns Allowed									
Type A	30	18	6	260	830	127	1250″	3461″	994″
Type B								4552	2040
Returns Prevented	25	14	12	204	573	220	1208	1772	736
Part									
Pure Part	10	13	11	107	409	214	538	1654	712
Progressive	14	9	11	96	148	126	536	1152	672
Direct Repetitive	24	12	15	120	237	155	716	1576	936
Reversed Repetitive	20	10	20	126	139	212	764	1297	1312

*Group averages to show the number of trials, errors and length of time required to learn certain problems by different "whole" and "part" methods. The first problem is the typical pencil maze. The A–Series is a list of thirty-two nonsensical syllables, the B–Series being the same syllables arranged in pairs. The criterion for having the maze learned is its perfect traversal; for the A–Series, successful anticipation of each successive syllable; for the B–Series, successful statement of the paired associate. The several "part" methods differ only as regards the manner of connecting the several parts. Of the "whole" methods Type A allows unrestricted trials upon the material and a measurement of the learning after each such trial; Type B allows the subject to completely learn the material in private before being tested, the sole learning measurement being in terms of time.

A comparative study of the data secured (See Table I) reveals three important conclusions. (1) So far as the whole and part methods are concerned, all the part methods prove superior to the whole methods, both for maze learning and for the highly comparable A–Series of nonsensical syllables, this superiority being demonstrated throughout by all the three criteria of learning, namely—trials, errors and time. Also, only the wasteful repetitive methods are inferior to the whole method when the B–Series is concerned, and this inferiority is only shown in the unavoidable requirements of trials and time, not in regard to errors. It suggests that less reviewing would have secured all the advantages of these repetitive methods and produced results uniformly superior to those of the whole methods. (2) One of the part methods, namely, the "progressive part," wherein each new section is learned as a unit and added at once to the earlier learned materials, is consistently most efficient throughout all the various learning tests. In maze learning, it is superior to its nearest competitor both in time and errors; in both the A–Series and B–Series, it is overwhelmingly superior in all three respects,—trials, errors and time. (3) Motor learning and learning verbatim obey, in the main, the same laws of learning, as witnessed by the facts that, in both cases, the part methods are more efficient than the whole; that the same part method is consistently the best in each field; that marked positive correlation invariably maintains between the results secured in the two fields by the entire list of methods, irrespective of whether the correlation relates to the trials, errors or learning time.

The conclusions suggest two points of explanatory value. It seems that any part method has certain inherent advantages, although these values may be lost sight of if the complex act of unit connection is not properly controlled. The advantages can be mentioned only in passing. The first of these is the matter of transfer. It is possible to show that the learners master successive units with increasing ease. The second advantage refers to the relation between learning effort and the length of the material. Neither for motor learning nor for this verbal material does the learning effort vary directly with the length of material. Diminishing returns are invariably secured as the material is lengthened. Any part method secures complete utilization of the items of transfer and avoids the diminishing returns due to the excessive length of the motor problem. Any whole method fails in both respects.

The second point of explanatory value refers to two principles operative in all learning. The first of these is that of *elimination.* It connotes the detection of the critical points in the problem, the careful study of all the details, the formation of proper associations, the rejection of others, etc. Consciousness is here at white-heat. The longer and more difficult the problem, the greater the task upon the learner to see the many details of the problem and to learn to eliminate the possible faulty reactions. Confusion, hesitation, emotional conditions all operate to delay the learning. If the problem is short, as is always the case with part methods, the principle of elimination can operate more effectively. The second principle is that of *mechanization.* The final stage of learning is no longer explorative and eliminative but rather, mechanizing and rendering habitual the entire activity. Whole method learning presents so many critical details that the principle of mechanization is not only delayed in being given an opportunity to operate but is repeatedly nullified by the reinjection of the highly conscious eliminative principle. Part methods utilize these principles to greater value. The explorative or eliminative principle is operative when the details of the short and easy parts are being mastered. The mechanizing principle then becomes operative, both logically and chronologically. It secures the mechanization of each unit and then the connection of the several units, being concerned only with rapidly welding the several unit habits. Hence any part method secures complete utilization of the principles of elimination and mechanization. Any whole method fails in both respects.

In conclusion, it may be said that the demonstrated advantages of these part methods when operating in the fields of motor learning and in learning nonsensical series verbatim, lead the writer to question the superiority of the whole method in the remaining field of meaningful material until the results of certain modified forms of part learning have been obtained herein and found inferior. This work is well under way.

THE RANGE OF INFORMATION TEST IN BIOLOGY II. ZOOLOGY*

N. M. GRIER

Central High School, St. Louis, Mo.

I. Introductory

One object in conducting investigations on range of information[2] in common-school branches has been to answer in part at least a question of economy in school administration—how much knowledge of the branches represented in both secondary and grade schools may we expect pupils to carry from the grade into the next higher division of the school system? With the formal barrier between the grades and the older type of secondary instruction broken down by the nascent junior high school and its characteristic general science, this question is important in connection with the content of biological science taught in secondary schools, since it is preceded in the grades by physiology and nature study. The matter of physiology has been taken up in the initial paper of this series,[3] (2) and it is proposed to treat in a similar manner the two main components of "nature study," elementary zoology and botany. The writer aimed also, in line with views of his urged in a previous note[4] to indicate how general science might be articulated with the semi-scientific branches of the grades so as to form in the school system a logically complete unit.

The problem of high-school zoology is mainly one of content. More disagreement exists here than in any other science taught in the high school. Two questions are at issue: First, To what extent shall "natural history" form a part of the course in zoology? If it can be shown that pupils entering from the grades are comparatively ignorant of those main forms of animal life which enter into direct economic relation with man, then high-school zoology must make provision for this neglect in nature study.

If, on the other hand, it may be shown that "natural history" is well taken care of by the grades, aided by the taste which preadolescent pupils have for reading of this type, the time customarily given it in many high-school courses of zoology may be devoted to a more intensive study of this science, such as usually characterizes

*The writer is indebted to the authorities of the St. Louis Schools, and Prof. L. M. Dougan, Principal Shaw School, St. Louis, for aid in the furtherance of this investigation.

high school physics and chemistry. Lloyd and Bigelow[5] remark that "secondary pupils do not undertake certain studies of living animals with the enthusiasm and earnestness which characterize nature studies in the grade schools; to the average secondary pupil studies of living animals are not part of serious science work," and further "it may well be doubted whether a course devoted exclusively to natural history is altogether best for secondary work." Very few textbooks are now in use which attempt high-school zoology on a natural history basis, but most have a liberal and perhaps unnecessary sprinkling of it.

The second question is whether the study of physiology and of physiological principles shall form part of the course in high-school zoology, since an essential feature of biological thought is that the activities of the human body be interpreted from the comparative standpoint. One short-lived effort to solve the problem of the content of high-school physiology resulted in a combination of zoology and physiology known as "comparative physiology." Some light may be thrown on this aspect by the preceding paper of this series.

II. Method

The 100 test-words were selected in the following manner. A list was first prepared of all those words pertaining more directly to botany and zoology discovered in the spellers and readers previously used by the pupils in their grade-school work. A similar list was made from Boy Scout and Camp Fire Girls publications. From these lists, and from a list prepared by Professor Dougan, who recently had charge of the revision of the nature study courses in the St. Louis schools, one hundred of the more frequently occuring words were checked. In selecting the words of his list the writer like Professor Dougan chose: (1) words in fairly common use, a knowledge of which would indicate the range of information in the natural history of animals of outstanding interest to man; (2) words, a knowledge of which would indicate some familiarity with physiological processes and which would be representative of lower animal groups as well as of man.

Since the method of giving the test has been discussed previously,[2] there is given here only the list of words used. I may explain only that in presenting the test, the pupils were told how the terms were selected and it was understood that they pertained directly or indirectly to animals.

III. Results

The results are found in the following tables. By "first-term students" is meant those taking up the study of botany at the time. "Second-term students" have finished botany and are beginning physiology. It may be well to remark that as a check on the original marking of the pupils ("first series"), the 100 words were defined by the examiner from the glossaries of text-books containing the words, and each pupil revised his own paper by placing a "second series" of marks after each word indicating the manner in which he should have marked it. A comparison of the D's, E's, F's and N's of the first and second series, besides showing fairly accurately the extent and nature of the error due to ignorance or misunderstanding of the real meanings, gives a rough preliminary measure of the extent to which these 100 terms are understood by first-year high-school pupils.

Test Words

abdomen	domestication	moth	rodent
absorption	down	nape	ruminate
antennae	drone	nestling	sail
auricle	drumming	nocturnal	scale
barb	excretion	non-retractile	shaft
beak	fin	nutrition	silkworm
beetle	gill	nymph	skeleton
bivalve	gizzard	ovipositor	soar
biped	grub	oxidation	spiracles
burrow	gullet	pancreas	stomach
butterfly	herbivorous	parasite	summer-resident
call	hibernation	pearl	tadpole
canine	incisor	perching	talon
carnivorous	insectivorous	permanent-resident	thorax
caterpillar	intestine	pigment	transient-resident
cerebrum	gland	prehensile	vein
chrysalis	kidney	primaries	venom
circulation	larvae	proteid	ventricle
clams	liver	pupa	vermin
cloven-footed	locomotion	quadruped	wading
cocoon	maggot	queen	wax
coral	mammal	quill	web-footed
coverts	migration	respiration	winter-resident
cross-pollination	mimicry	reproduction	worker
diurnal	molars	retractile	yolk

TABLE I

Dependence of Range of Information on Academic Status

(1st series)

Academic Status	No.	D.	E.	F.	N.
1st-term pupils.........	26	14	27	15	43
2nd-term pupils........	87	19	30	22	29

(2nd series)

Academic Status	No.	D.	E.	F.	N.
1st-term pupils.... ...	26	23	29	12	35
2nd-term pupils........	87	25	30	19	26

While it is evident from Table I that with maturity there is some increase in the number of terms that can be defined (D), explained (E), and are familiar (F), with a corresponding decrease in the number of new terms (N), yet as was shown also in the table on academic status in the preceding paper in the series, the greatest change is in the number of F's. The difference in D's and E's is so slight in both series and decreases to such an extent in the second that so far as a knowledge of "Natural History" is concerned, the writer is inclined to attribute the increase to the more highly organized motor processes of the second-term pupils, as shown by their ability to carry out more readily the writer's instructions while giving the test. This is somewhat substantiated by the fact that only 3 of the test words used occur in the preceding course of botany. The checking up process increases the number of D's and E's, and decreases the number of F's and N's in this table as well as the others. Some confirmation of Lloyd and Bigelow's statement that the pupil's interest in natural history declines at this age is shown by the want of a more decided increase in the number of D's and E's, also by the average error of marking, 6 per cent. in 1st-term pupils, 8 per cent. in second-term pupils. If their interest was a more active one, a greater percentage of D's and E's could be expected in second-term pupils, and they might be expected to show a less percentage of error. At best it seems that discriminatory power in the subject of Natural History has not advanced.

These results reclassified by sex do not show significant differences between the sexes, except in the percentages of error after second checking. These indicated that the boys were more accurate in judging the exactness of their comprehension of the 100 terms than were the girls.

A list of typical errors of definition follows, in which the words that apparently are the source of confusion are indicated in parentheses. Comparison with the list published in the previous article

of this series will show how deep-seated the errors in the meaning of terms may be, and, as previously suggested, will furnish opportunities for objective attacks by teachers in solving the first problem indicated in this paper. It is felt that they form a basis, conscious effort with which on the part of the teacher will serve to put the subject matter of nature study into such form as will readily carry over into zoology.

Antennae—sting of a bee: feelers on a butterfly: part with which an insect sees and feels, and like a long bar.

Abdomen—stomach: lower part of bowels.

Absorption—taking in of liquids in form of vapor: taking in of oxygen.

Auricle—vein through which blood passes from the heart.

Butterfly—a fly with large wings and body: a little worm that has wings and flies: a moth like a bug brightly colored.

Beetle—a small bug: quadruped creature: animal derived from fly family.

Biped—an object built on a plan of 2 (Bilateral Symmetry).

Beak—nose piece of an insect (proboscis): mouth of a duck (bill).

Burrow—a species of mule used mostly in the mountains (burro).

Bivalve—double value (bi-valve).

Caterpillar—a worm which crawls on flowers and on ground: is a silk worm: inside of a butterfly.

Cocoon—a little winter home for a silk worm.

Carnivorous—that which eats human substance.

Clovenfooted—the skin between the toes: to have feet like a horse.

Coral—stockade of animals (corral): skeleton of a fish: parasite animal of the sea: a gathering of dead animal matter from the ocean.

Canine—pertains to dogs, especially teeth: a poison, some in coffee (caffein).

Drone—a bee that is lazy and lays eggs: a female which does not work.

Drumming—a man who goes around selling goods.

Down—below.

Fin—a sharp pointed projection on the back of the fish used for protection.

Gill—¼ pint.

Grub—rough way of saying "food."

Incisor—to insert or drill.

Intestine—soft part of stomach: eggs of moth.

Kidney—bone on inside of stomach.

Larva—hot substance given off in an eruption (lava).

Nymph—a water object consisting of a human body and fish.

Proteid—sugar or starch food.

Scale—thorny substance on the back of a fish.

Yolk—white of an egg.

---

[2]Whipple, G. M., *A Range of Information Test*, Psych. Review, 16: 1909, 347–51; also *Manual of Mental and Physical Tests*, Test 51.

[3]Grier, N. M., *The Range of Information Test in Biology; I Physiology*, Vol. 9: 1918, p. 210, *This Journal.*

[4]Grier, N. M., *Huxley and General Science*, School and Society, Vol. 6, No. 136.

[5]Lloyd and Bigelow, *The Teaching of Biology.*

THE INFLUENCE OF PRACTICE ON THE CORRELATION OF ABILITIES

GEORGINA I. STICKLAND

Barnard College, Columbia University

Correlations which have been computed between mental tests have been, for the most part, comparisons of initial trials or of records made after a few preliminary trials. Only in one case was an experiment performed in which it was possible to correlate the abilities of individuals who had in all probability reached their practice limit in the tests used. Professor Hollingworth reports an investigation in which he compares records made after one, five, twenty-five, eighty, and two hundred and five repetitions of each of seven tests.* He finds a high correlation between some of these tests (*i.e.*, the familiar adding, tapping, opposites, color-naming tests) after the two hundred and fifth trial. He finds a progressive increase in the average of correlations of all tests from the first to the two hundred and fifth trial. This increase of co-efficients with practice and final high correlation leads the experimentor to favor some form of the doctrine of "general ability," to suppose that there may be, when a practice level is reached, a positive correlation of desirable traits.

The experiment which follows was prompted by a desire to test this hypothesis further. Five tests were used in the investigation. These were:

Color-naming: Two Woodworth-Wells blanks pasted together. Record, time required to name correctly a set of two hundred colors. Test repeated twenty-five times.

Tapping: Record, number of taps executed in two minutes with the hand stylus, right hand. Test repeated twenty-five times.

Adding: Kraepelin blank used. Record, time required to add one hundred examples of two numbers each. Test repeated thirty times.

Multiplying: Record, time required to solve mentally a set of five multiplication examples. The examples required the multiplying of one two-place by another two-place number. No number below four was used—no example was repeated. Test repeated twenty-two times.

---

*H. L. HOLLINGWORTH, *Vocational Psychology*, Ch. xi; also *Correlation of Abilities as Affected by Practice*, J. of Ed. Psych., Sept. 1913, pp. 405–414.

Word-building: Record, number of shorter words built from a given word (as weather, psychiatry, etc.) in a specified time. Words varied in difficulty and the amount of time was changed as the experiment proceeded. The test had twenty-two repetitions.

It is important to note the difference between these last three tests and those used in Professor Hollingworth's investigation. A repetition of the tests used in the former experiment meant a going-over of identical material. Definite bonds were formed and perfected. Responses in the opposites and adding tests became like responses in the color-naming test. The adding, multiplying, and word-building tests of the present investigation were chosen with a view to eliminating, as far as possible, the probability of a change in the nature of the tests. Responses cannot be stereotyped for they are changed at each repetition. In the adding experiment, there were six different sets of one hundred numbers each used. There was no possibility, then, of learning the order of appearance of the answers. In the multiplication test we have each time a different problem, and in the word-building a different word. Practice in these cases means not so much the strengthening of a few particular bonds, as acquiring experience in the exercise of some "mental function."

The subjects in the experiment were fifteen college students. Each subject spent two separate half-hours a week in practice. The entire experiment extended over a period of from seven to eight weeks. All records were taken, with a stop watch, by one person.

In order to find the relation of an individual's proficiency in one sort of test to her proficiency in the others at different points of the curve of practice, the practice period was divided into eight parts, at each of which parts the records were correlated. These divisions were:

1st Division. The initial trial.

2nd Division. The second, third, and fourth trial of color-naming, tapping, multiplying, word-building. The second, third, fourth and fifth trial of adding.*

3rd Division. Fifth, sixth, seventh of four tests. Sixth, seventh, eighth, ninth of adding.

4th Division. Eighth, ninth, tenth of four tests. Tenth, eleventh, twelfth, thirteenth of adding.

5th Division. Eleventh, twelfth, thirteenth of four tests. Fourteenth, fifteenth, sixteenth, seventeenth of adding.

*The adding test was the shortest and was given the largest number of trials. Therefore four adding trials were united to make one division, while only three of each of the other tests were used.

6th Division. Fourteenth, fifteenth, sixteenth of four tests. Eighteenth, nineteenth, twentieth, twenty-first of adding.

7th Division. Seventeenth, eighteenth, nineteenth of four tests. Twenty-second, twenty-third, twenty-fourth, twenty-fifth of adding.

8th Division. Twentieth, twenty-first, twenty-second of four tests. Twenty-sixth, twenty-seventh, twenty-eighth, twenty-ninth of adding.

At each of these points the average record made by each individual in each test was correlated with her average record in other tests. The correlation formula

$$R = 1 - \frac{6\Sigma D^2}{n(n^2-1)} \text{ was used.}$$

The details of the correlation are omitted in this article. Table I gives the average correlation of each test with all other tests, at eight points in the practice period. The coefficients in the original table for each test at each point, have been averaged to make Table I.

TABLE I

Average Correlation of Each Test With All Other Tests, including Multiplying.

Average Used as Measure

Divisions	1	2	3	4	5	6	7	8
Color-naming......	.14	.15	.26	.28	.33	.32	.24	.31
Tapping..........	.22	.22	.28	.19	.29	.38	.12	.16
Adding...........	.23	.19	.43	.36	.45	.48	.42	.35
Multiplying........	.17	.06	.19	.18	.20	.25	.12	—.10
Word-building......	.24	.52	.27	.34	.42	.45	.39	.26
Average...........	.20	.23	.29	.27	.34	.38	.26	.19
Average of each two	.215		.28		.36		.225	

Here a gradual increase in coefficients is noticeable which reaches its maximum at the sixth division, where the average correlation of all tests with all other tests is +.38. After this point there is a decrease which is slight in all tests, but multiplying.

Further inspection of the table shows that the multiplying test is characterized by low and irregular correlation with all other tests but adding. The multiplying test, the most complex of all, was the one in which there was the least approximation to a practice limit. The fact that in this test there was not enough practice to eliminate variability or to equalize the effects of previous practice might account for this lack of definite correlation. It is noticeable that the correlation did increase up to the sixth division, where it became +.25, but after that, through accidental variation or other causes, it became low again.

TABLE II

Average Correlation of Each Test with all other Tests, excluding Multiplying. Average Used as Measure

Divisions	1	2	3	4	5	6	7	8
Color-naming......	.16	.32	.36	.41	.46	.45	.34	.44
Tapping..........	.28	.25	.28	.22	.39	.49	.26	.30
Adding...........	.20	.17	.40	.30	.36	.50	.45	.45
Word-building.....	.26	.66	.33	.41	.50	.42	.38	.38
Average..........	.23	.35	.34	.34	.43	.49	.38	.39
Average of two.....	.29		.34		.46		.385	

TABLE III

Average Correlation of Each Test with all other Tests, excluding Multiplying. Median used as Measure

Divisions	1	2	3	4	5	6	7	8
Color-naming......	.16	.32	.38	.42	.46	.47	.35	.31
Tapping...........	.28	+.06	.26	.19	.43	.43	.18	.17
Adding............	.20	−.04	.45	.24	.49	.49	.33	.33
Word-building.....	.29	.27	.32	.41	.47	.45	.32	.26
Average...........	.23	.15	.35	.32	.36	.36	.30	.27
Average of two.....	.19		.335		.36		.295	

In order to eliminate the tendency of multiplying to lower all coefficients, other tables were prepared from which correlations with multiplying were omitted. Table II shows the average correlation of each test with all other tests excluding multiplying. There the coefficients are higher and their upward trend even more obvious. They increase steadily up to the sixth division where they fall off a bit, never becoming, however, as low as they were in the first, second, third and fourth divisions. Table III shows a like result. The difference between this table and the two preceding ones is that here the median record of each division was used as the unit of measurement, whereas in Table I and II the records in each division were averaged.

TABLE IV

Average Correlation of each Test with all other Tests, excluding Multiplication. Best Record Used as Measure

Divisions	1	2	3	4	5	6	7	8
Color-naming........	.16	.29	.55	.45	.51	.54	.50	.44
Tapping.............	.28	.28	.25	.19	.42	.51	.26	.34
Adding..............	.20	.22	.46	.24	.47	.57	.45	.41
Word-building........	.29	.33	.40	.32	.54	.61	.34	.53
Average.............	.23	.28	.42	.30	.49	.56	.39	.43
Average of two.......	.26		.385		52.5		.41	

In Table IV only the individuals' best trials in each division were considered. Accidental causes, of poor records were eliminated here. One would expect, if correlations do increase with approach to a practice level, that where highest ability only was measured, coefficients would be higher than they are in the tables computed from average and median measurements. This is the case. All average coefficients are higher. The best coefficient ever reached (56 per cent. in the sixth division) is found here.

TABLE V

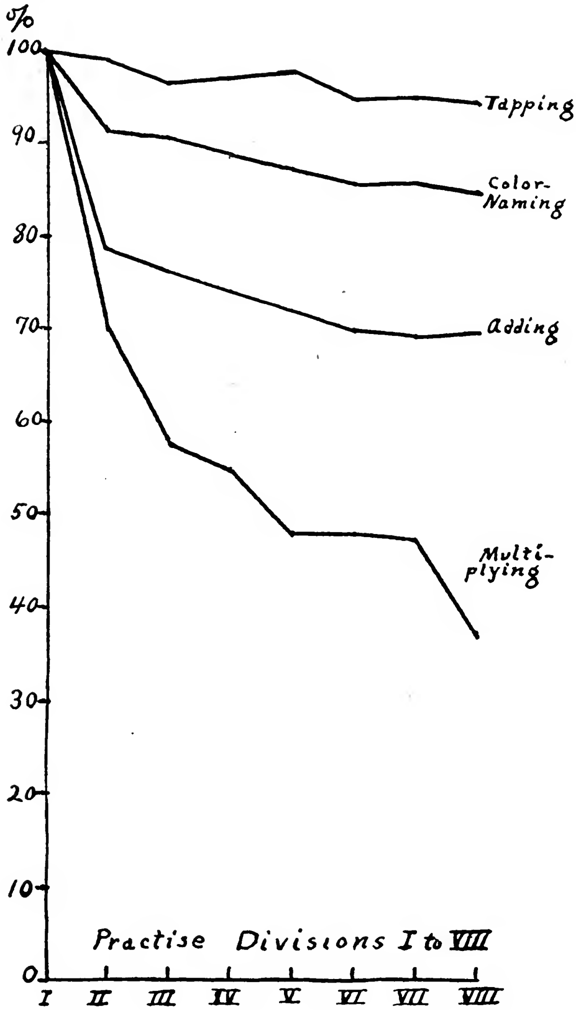
Correlation of each Test with every other excluding Multiplication at Highest Point Reached by Each Subject

	Color-naming	Tapping	Adding	Word-building
Color-naming..........	...	.49	.53	.36
Tapping..............	.49	...	.21	.37
Adding...............	.53	.21	...	.52
Word-building.........	.36	.37	.52	...
Average..............	.46	.36	.42	.42
Average of Average.....		.42		

This is brought out even more clearly by Table V. The eight divisions are done away with, and each subject's highest point only is considered. This is the point which might have become the average if sufficient practice had been given. The correlation is relatively high—the average is 42 per cent.

Tables II, III, IV and V all show an increase in correlation up to the sixth point, followed by a slight decrease which makes the last two coefficients, though greater than those of the first to the fourth periods, a little less than the coefficients of the fifth and sixth divisions.

Individual practice curves, and averages of these curves show this same phenomenon. Inspection of the individual curves (not shown in the article) reveals great improvement up to the sixth division, followed by relatively little or entire lack of improvement. In the figure the records of the fifteen individuals have been averaged. The amount of time required for the initial trial of each test has been taken as the standard and averages of subsequent practice divisions are represented in terms of per cent. of this record. No word-building curve is shown, for the varying difficulty of the words and the different amounts of time allowed make such a curve worthless. The curves of the other four tests show gradual, steady decrease until the sixth division, after which there is an average


%
100
90
80
70
60
50
40
30
20
10
0
Tapping
Color-Naming
Adding
Multi-plying
Practise Divisions I to VIII
I
II
III
IV
V
VI
VII
VIII

improvement of only seven tenths of one per cent., in color-naming, five-tenths of one per cent., in adding, and three-tenths of one per cent. in tapping. Multiplying shows a greater improvement (10 per cent.) but even this is small when we compare it with the 52 per cent. improvement gained before this. It seems as though at the sixth division some sort of a practice level is normally reached. This point corresponds exactly with the place where increase of correlation coefficients has ceased.

This correspondence might imply a casual relation of some sort. The lack of improvement in ability might be suggested as the cause of lack of improvement in inter-correlations. *Where practice improves performance, correlations increase.* Where there is a fall in effort, ability, or what-not, or where accidental causes lower proficiency, the coefficients stand still or show a slight decrease.

The results of this experiment seem to show a positive correlation increasing with practice between the abilities measured by these tests. This is brought out by a correlation of median points and average points of all tests but multiplying, confirmed by a comparison of best records, that is, of records representing more closely the approach to the practice limit. The fact that multiplying, a more complex, less easily mastered test, correlates poorly need not weigh against our hypothesis, but may even be urged in support of the assumption that higher proficiency is needed for greater positive correlation. A comparison of practice curves with correlation coefficients brings further evidence for the theory that improvement in ability is a factor in the increase of correlations.

The great similarity of the subjects used and the dissimilarity of the tests makes these results fairly convincing. But the small number of subjects available gives opportunity for many accidental errors, which further experimentation, now being carried on along this same line, may eliminate. The point here considered relates to one only of a number of problems which have practical and theoretical importance and which justify a detailed inquiry into the psychology of practice limits.

The Journal of Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD.

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

NOTICE TO SUBSCRIBERS. Notify the publishers, Warwick & York, Inc., of any change of address; as copies lost in transit because of subscriber's removal to another address will not be duplicated.

THE JOURNAL is issuing a series of Monograph Supplements for the publicaiion of more extensive researches in educational psychology. For details prospective contributors may apply to H. E. Buchholz, 10 East Center St., Baltimore, Md.

The price of the JOURNAL is $3 00 per year of ten monthly issues, payable in advance. Single issues are 50 cents each. Back volumes are $5.00 in original wrappers and $6.25 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME IX SEPTEMBER, 1918 NUMBER 7

EDITORIAL

Personnel

Never before in war has such explicit and varied use been made of psychology. From the mental tests given to all enlisted men to the special and highly involved studies in gun pointing, range finding, and aviation, psychologists are being called upon to assist the government to utilize to the utmost the brain power of the country. The slogan in every department is "the right man for the right place." An important factor in the realization of this ideal has been the work of the Committee on Classification of Personnel in the Army under the direction of Dr. Walter Dill Scott. To further the work of this committee a new periodical has recently been established by the Adjutant General entitled *Personnel*. The object of this periodical is to bring into closer touch the various agencies and officers that have to do with the selection of men for special tasks, and to enable the Adjutant General and other government officials to reach those with whom they wish to communicate more quickly and effectively.

The first number of the new journal contains the address of the Secretary of War, Newton D. Baker, before the eighth school for personnel adjutants at Camp Meigs, Washington. In the course of his address Secretary Baker gave a clear indication of the reaction which the mere term "psychology" calls forth in the average business or professional man.

"If one were to select, for instance, a general to command an expedition, he would be quite likely to select a man whom he liked as a man, and not with any relation whatever to his capacity to command the expedition. And this is true, generally, so that some system of talents which is not affected by immaterial principles or virtues, no matter how splendid, something more scientific than the haphazard choice of man, something more systematic than preference or first impression, is necessary to be devised. When Dr. Scott presented to the War Department the question of calling in psychology in this work we all realized that that science had a particularly unfortunate name; it sounds very highfaluting and cloud-stepping as it were. It is the skyscraper, so far as its name is concerned, of modern science, and there is a certain revulsion in the ordinary man's mind when you speak to him about a thing that sounds as pretentious as that. I commend to my good friend Dr. Scott that he try to see whether they can't change that term to the study of human action or something with a popular appeal. I am sure the impediment would be so much less if that could be done."

Now why should the mere term psychology be such an impediment? Why should it produce such a revulsion in the ordinary man's mind? In itself it is no more formidable than aerodynamics or ballistics. It must be on account of the associations that surround the word in the lay mind. Is it on account of its traditional association with metaphysical philosophy that psychology is regarded with suspicion, or is it because through newspaper articles, novels and other sources the term has been popularly connected with the study of spooks? At any rate the passing remark of the Secretary is significant of the need for popular enlightenment as to the scientific use of the term.

But the Secretary soon found that psychology, as represented by the Committee on Personnel, had certain very definite and specific things to offer, and he came to the conclusion that these things might be of distinct service in winning the war.

"The problem is to weld those millions not merely into an army which will fight —any American army will fight, we have discovered that; not merely into an army which is willing to die, if necessary, in order to maintain its position and uphold its traditions—we have discovered now that this is true of Americans everywhere; but an army which will fight with the least wastage, the most effective execution, the most intelligent co-operation, the most complete because the most understanding subordination, and that sort of an army comes only when men are doing the things for which they are by nature and training best adapted. Now that is a very high task, it is a task of contributing the finishing touches of efficiency to this great American army. I have myself been tremendously stimulated, not only by the care and wisdom employed in assigning each enlisted man but also by the result which Dr. Scott and his associates have obtained in the officer selecting tests which they have devised. The Rating Scale has been enthusiastically received by the Army, but I am told officers have occasionally objected to answering some of the Binet test questions the doctor has put to them; I have been rather afraid to take a Binet test myself, but when you finally come down to it, the Rating Scale and tests for mental alertness are the application of a perfectly rational method to the great problem of putting a man in the position where he can be of the most service to the country and to the common cause."

Thus, in spite of his aversion to the term, the Secretary pays high tribute to the service which psychology is rendering in perfecting the greatest fighting machine which this country ever contemplated.

J. C. B.

NOTES AND NEWS

The similarity between the Otis Absolute Point Scale for the Group Measurement of Intelligence, published in the May and June issues of this JOURNAL, and the military group intelligence scale used in the examination of recruits has been to some a source of wonder. It is fitting, therefore, that a statement be made regarding the priority relationships of the two scales. The Otis Scale was originally drafted in 1916, and, according to Professor Lewis M. Terman, a member of the committee, Mr. Otis generously placed his material at the disposal of the Committee on Psychological Examination of Recruits, which drew up the military group intelligence scale. The ingenious arrangement of the Otis tests for group examining served largely as the basis for the intelligence scale adopted by the committee for army use.

The work of educational reconstruction for wounded soldiers is making good progress at the Walter Reed General Hospital, near Washington. In July there were 488 soldiers taking one or more courses in the curative workshops and bedside occupational work at the hospital. Over thirty courses are offered, ranging from shoe repairing to commerical law, and five new buildings have recently been opened. The methods and materials of each course are adapted to the patient's particular needs or handicaps. The work occupies the time of fifty instructors, and is in charge of Major B. T. Baldwin, educational director and chief psychologist.

The Wisconsin Anti-Tuberculosis Association announces at the University of Wisconsin a course for health instructors, designed to train in a ten months' period a limited number of women in the fundamental knowledge of health conservation and disease prevention. The teacher with her professional training and her experience in dealing with children seems especially qualified to become this new type of worker—the health teacher. There is need for the health teacher in schools, in city health departments, in factory welfare departments, in health stations and clinics, and in many other of the reconstruction activities which America must undertake on an increasingly large scale as the war continues. The work is under the direction of Dr. E. V. Brumbaugh, and includes courses in fundamental principles of health administration, social service, rural health inspection, first aid, health games and corrective exercises, and health propaganda.

In accordance with a proclamation of President Wilson a campaign for infant conservation is being conducted by the health committee of the mayor's committee on national defense of New York City. The campaign is in charge of Dr. Josephine Baker,

director of the bureau of child hygiene in the health department, and almost fifteen hundred physicians, nurses, welfare and social workers have volunteered their assistance. It is estimated that there are over 600,000 children to be examined.

Dr. Frank V. Thompson, assistant superintendent of schools, has been elected superintendent of schools in Boston. Dr. Thompson was formerly principal of the Boston High School of Commerce, and has been prominently identified with the furtherance of vocational education in the United States. It is a noteworthy indication of the trend of thought in education that the new superintendents of both New York and Boston are men who are conspicuous for their interest in vocational education. It is to be hoped that this points to the reorganization of our school work on the basis of greater efficiency.

Dr. Walter S. Monroe, director of the bureau of educational measurements and standards at the Kansas State Normal School, Emporia, has accepted the position of director of the bureau of co-operative research at the University of Indiana.

Dr. Arthur Holmes, dean of the general faculty of the Pennsylvania State College, and author of *The Conservation of the Child*, *Backward Children*, and other works, has been elected president of Drake University, Des Moines, Iowa.

In their service bulletin for September *The World Book*, published by Hanson-Roach-Fowler Company of Chicago, present five interesting prize questions for the month, a reference outline of how our country's laws are made, the story of the month, and seventy "look-it-up" questions for the World Book Club. Teachers will find these bulletins very suggestive.

An important book by Professor James Burt Miner, of the Carnegie Institute of Technology, entitled *Deficiency and Delinquency; An Interpretation of Mental Testing*, is announced for publication in September by Warwick and York, Baltimore. The book presents a critical analysis of the numerous studies of defective delinquents, and contains much new material resulting from original investigations. It is one of the most thorough and fundamental discussions of the important question of mental and moral borderlines that has yet been produced.

Another book announced by Warwick and York is *Graphology and the Psychology of Handwriting*, by Professor June E. Downey, of the University of Wyoming. Psychologists are familiar with Professor Downey's investigations of handwriting, and will welcome this systematic consideration of the entire subject.

A significant book on *The Psychology of Childhood,"* by Naomi Norsworthy and Mary Theodora Whitley, of Teachers College, Columbia University, has just been issued by Macmillan, New York. An original feature of the book is the attempt to present a cross section of child life at the ages of five and eleven.

One of the most widely discussed war books is *Under Fire,* by the French artist and literary genius Henri Barbusse (E. P. Dutton, New York). Hotly attacked as too realistic and gruesome by some, it is vigorously defended and highly praised by the great French dramatist Edmond Rostand, who calls it a great poem, tumultuously and admirably arranged.

In *The Present Conflict of Ideals* Professor Ralph Barton Perry, of Harvard University, makes a study of the philosophical background of the world war. The author endeavors to relate the conflicting national ideals of the war with the philosophical tendencies of naturalism, idealism, pragmatism, and realism. The book is promised by Longmans, Green and Company for the last of September.

Psychic Tendencies of Today by Alfred Martin is one of the new fall books offered by Appleton, New York. Spiritualism, psychic research, theosophy, Christian Science, and New Thought, are discussed, and Sir Oliver Lodge's book *Raymond* is analyzed. The author believes that modern materialism, instead of denying immortality, shows the scientific necessity for belief in it.

The rapidly increasing share which women are called upon to take in the political life of the country makes *The Woman Citizen,* another of Appleton's September publications, very timely. The author is Professor Horace A. Hollister, state high school visitor at the University of Illinois, and the book is a survey of woman's status and achievement in the past and the present, and a discussion of the needed readjustments of woman's place in society.

PUBLICATIONS RECEIVED

LEONARD P. AYRES. *A Measuring Scale for Handwriting.* New York: Russell Sage Foundation, 1917. Five cents.

This scale is a revision of the author's 1912 "Three Slant Edition," and since all the samples are from the first lines of Lincoln's Gettysburg Address, it is named "The Gettysburg Edition." Directions are given for securing samples, grading them, and comparing with standard distributions of both quality and speed.

FRANK W. BALLOU. *Arithmetic. The Value to the Teacher, to the Principal, and to the Superintendent of Individual and Class Records from Standard Tests.* Boston Public Schools, Bulletin No. 13, of the Department of Educational Investigation and Measurement, 1917. Pp. 83. Seven cents.

An important practical contribution to the use of educational measurements in the class-room and in school administration. Suggestions are given for the interpretation of representative graphs, the utilization of records by the teacher, the eradication of deficiencies, the treatment of cases of marked accuracy and instances of exceptionally high and low ability.

DR. H. BERNHEIM. *Automatisme et Suggestion.* Paris: Libraire Felix Alcan, 1917. Pp. xv, 168. Fr. 2. 75.

Perhaps no one has done more to reduce the phenomena of hypnotism to a reasonable basis than Bernheim, of Nancy. The present book represents a quarter century of study and discussion on the subject of suggestion, and gives the ripely matured views of the author. Among the questions considered are the nature of reflex automatisms, automatism in mental life, hypnotism and artificial somnambulism, freedom and moral responsibility, psychoneuroses and psychotherapy. The author's views are entirely in accord with the behavioristic conception of mental life.

CHARLES SCOTT BERRY. *A Study of Retardation, Acceleration, Elimination, and Repetition in the Public Elementary Schools of Two Hundred Twenty-five Towns and Cities of Michigan.* (Reprinted in part from the Seventy-ninth Annual Report of the Superintendent of Public Instruction of the State of Michigan for the Year 1915–1916. Pp. 47.)

An elaborate statistical study. In no case do the accelerated pupils balance the retarded pupils.

HENRY BROWNE. *Essays on the Reform and Revival of Classical Studies.* New York: Longmans, Green and Company, (No date). Pp. xvi, 281. $2.60.

This eminent Jesuit classicist, writing from the point of view of the present educational crisis in England, realizes the change which the war is forcing upon educational thought and practice, and meets the issue very frankly. He is, however, an ardent advocate of reform in classical teaching, and believes that the stirring of educational ideas will serve to eliminate many of the present drawbacks to this work, and will revivify the classical spirit. He is espcially keen on the utilization of archeology, numismatics, and other phases of modern research in awakening in the minds of pupils an appreciation of the richness of Greek and Roman civilization.

H. Addington Bruce. *Handicaps of Childhood.* New York: Dodd, Mead and Company, 1917. Pp. ix, 310. $1.50.

The author argues "that in view of the discoveries of modern psychology with regard to individual development the mental and moral training of children by their parents ought to be begun earlier, and be carried on more intensively, than is the rule at present." Hence the book sums up in easy, narrative form those suggestions from modern psychology and education that will be helpful to this end. Some of the handicaps considered are mental backwardness, being the only child, a sulking disposition, jealousy, selfishness, bashfulness, indecision, stammering, and a too vivid imagination.

Bureau of Educational Experiments. Playthings. Study of Animal Families in Schools. Experimental Schools. A Bibliography of Psychological Tests. New York: Bureau of Educational Experiments, 1917. Bulletins 1–5, ten cents each. Bulletin No. 6 (Bibliography) twenty-five cents.

These booklets contain excellent suggestions about playthings and animals in the schoolroom, describe several interesting private schools, and present a selected bibliography of articles on mental tests.

K. M. Dallenbach. *Blindfold Chess: The Single Game.* Reprinted from Studies in Psychology, Titchener Commemorative Volume, 214–230.

A masterly analysis of the thought activities involved in this most intricate of reasoning games, and a detailed description of the controls employed when the customary visual control was removed.

John Davidson. *Means and Methods in the Religious Education of the Young.* New York: Longmans, Green and Company, 1917. Pp. viii, 152. $1.00.

This book was written as a guide for Sunday School teachers and emphasizes the need for teaching religion, habits of right thinking and feeling, the treatment of the miraculous, the parables, story-telling, aims and modes of questioning, and the use of pictorial illustrations. The procedures recommended are frequently illustrated by examples.

Educational Survey of the Public Schools of Brookline, Mass. Published by the School Committee, 1917. Pp. vii, 436.

The survey briefly reports the application of the following tests: The Courtis Arithmetic Test, Series B; the Stone Reasoning Test; spelling tests from the Ayres scale and the Boston lists; the Ayres handwriting scale; the Holmes reading tests; and the Harvard-Newton composition scale. A demonstration school is urged. which shall combine the functions of a model and an experimental school.

A. Caswell Ellis. *The Money Value of Education.* Washington: Bureau of Education, Bulletin, 1917, No. 22. Pp. 52. Fifteen cents.

This is a striking presentation of the value of education to the nation, to the community, and to the individual. The study is graphically illustrated, and contains a bibliography of seven pages.

J. Berg Eisenwein and Marietta Stockard. *Children's Stories and How to Tell Them.* Springfield, Mass.: Home Correspondence School, 1917. Pp. xiv. 352. $1.50.

This valuable manual for teachers contains a discussion of the place of stories in child life, the selection and structure of stories, methods of story telling, inventing stories and teaching children how to invent them, and fifty stories to tell to children.

J. Rogues de Fursac. *Manuel de Psychiatrie.* Paris: Librairie Felix Alcan, 1917. Pp. vii, 509. Fr. 7.

This is the fifth edition of this celebrated *Manuel,* and in its revised and augmented form it will probably see as many years of usefulness as it has in the past. The first part of the book is devoted to general psychiatry, and deals with the causes, symptoms and treatment of mental disturbances. The second part treats of specific psychoses.

Arthur I. Gates. *Recitation as a Factor in Memorizing.* Archives of Psychology, No. 40, September, 1917. Pp. 104. $1.00.

An important study of the value of learning by reading as compared with learning by recitation in the case of elementary school pupils and university students. The first twenty pages review the work of previous investigators, the next forty pages describe the experiments and give the results, and the last forty pages present an analysis of reading and recitation as factors in learning. The best procedure was found to be to devote from twenty to forty per cent. of the time to reading and the remainder to recitation.

M. Annie Grace, Emma C. Monroe, and others. *Lessons Plans in Fourth Grade History.* Baltimore: Warwick and York, 1917. Pp. 155. Seventy-five cents.

Detailed lesson plans for a wide range of historical topics from the Trojan War in Asia to the French and Indian War in America. This book not only presents history, but shows just how to treat the topics to produce the most effect with young pupils. A mine of helpful suggestions for the elementary teacher.

F. M. Gregg. *Hygiene as Nature Study.* Peru, Nebraska: Published by the author, 1917. Pp. 170. Seventy cents.

This admirable little compendium prrovidee in compact form just those things that the teacher needs to present to a class of V to VIII grade pupils. There is no platitudinous moralizing, but the attack is strictly scientific. A feature of the book are the numerous drawings for blackboard sketches. Many simple experiments suited to elementary grades are described.

Percy Gardner. *Evolution in Christian Doctrine.* New York: G. P. Putnam's Sons. 1918. Pp. xiii, 241.

The author claims to stand as a connecting link between the old broad church and the new modernism. He frankly accepts evolution in science and critical method in history, and asks that the fundamentals of the Christian religion be examined afresh and restated in conformity to the intellectual requirements of the age. Especially does he feel the need for raising the intellectual standard of the clergy. In this book he traces the evolution of Christianity, and considers from the evolutionary point of view such doctrines as miracle, prayer, the immanence of Christ, the Holy Spirit, the future life, and other points of the creed. While scarcely radical enough to satisfy the thorough-going social psychologist, the book is a goodly step in the right direction.

A. J. Grant. *A History of Europe.* New York: Longmans, Green and Company, 1917. Pp. xiv, 778. $2.75.

This is a revised edition of a well-known and highly esteemed text-book on European history. It is to a large extent a political and social history, and thee narrativ

moves with a vigor and directness that enthralls the attention and captivates the imagination. Of especial interest are the chapters on the development of Europe since 1870, the Balkan wars, and the causes of the present war of nations.

JOHN RICHARD GREEN. *A Short History of the English People.* Revised and Enlarged, with Epilogue to 1914, by Alice Stopford Green. New York: The American Book Company, 1918. Pp. iv, 1039. $2.00.

This new and enlarged edition of Green's "English People" is particularly timely, and should find a cordial reception. The clarity of Green's style and the breadth of his interests have long been recognized, and have given his work its well-merited preeminence among histories of England. In the present edition the epilogue, by his daughter, occupies almost 200 pages and brings the narrative from 1815 down to 1914. We recommend the book most highly to our readers.

ALBERT BUSHNELL HART. *School History of the United States.* New York: The American Book Company, 1918. Pp. 505, xxxiv. $1.20.

In this admirable text for the upper elementary grades the author has endeavored to bring out the European background of our history, to give due consideration to the history of different parts of the country, to stress the social and industrial development of the country, and to give a clear idea of the way our government is carried on. A third of the book is devoted to the period since the civil war, and a special effort is made to familiarize boys and girls with problems which they will be called upon to face. The book is a magnificent example of the printer's art, containing excellent maps, many full-page illustrations in color, and a very great number of small illustrative insets.

WALTER MORRIS HART, Editor. *English Popular Ballads.* New York: Scott, Foresman and Company, 1916. Pp. 370. Forty cents.

The real spirit of a people is perhaps best seen in its ballads, and this little book renders valuable service in making these English ballads accessible to the many. A helpful feature of the book is the forty-page introduction.

C. GASQUOINE HARTLEY. *Motherhood, and the Relationship of the Sexes.* New York: Dodd, Mead and Company, 1917. Pp. 402. $2.50.

The relation of the sexes is becoming one of the greatest social and political problems of the day. Six millions of women have just been enfranchised in England; the suffrage amendment is being debated in America; Norway, Finland and Australia have led the way in the suffrage movement; large concessions to women are in sight in France, Russia, and Hungary; and even in Germany three powerful women's organizations are starting an agitation for suffrage, in spite of the imperial command not to raise this question until after the war. This political recognition is merely the reflection of the economic and social evolution that has been going on for a long time. But the readjustment of the sexes brings forward with new insistence some of the oldest and most puzzling social problems that have vexed human minds. These problems find clear, scholarly and convincing treatment in the present book. The author takes the biological method of attack, and traces parenthood in insects, reptiles, fishes, birds, and mammals, showing the increasing significance of maternal care. A section on the mother in the primitive family leads to the consideration of the modern woman, the family, the home, marriage, sexual relationships outside of marriage, the unmarried mother, and the necessity

for sex education. The three chapters on the latter subject are singularly wise, sane, and well-balanced. The book is a noteworthy contribution to a vital topic.

JOEL HATHAWAY. *Modern French Stories.* New York: The American Book Company, 1916. Pp. 220.

As a result of the war there will doubtless be a marked increase of interest in the study of French in the schools, and these charming stories will aid greatly in developing vocabulary and feeling for good French idiom.

FREDERICK HOUGHTON. *First Lessons in English for Foreigners in Evening Schools.* Pp. 150. *Second Book in English for Foreigners in Evening Schools.* Pp. 180. New York: The American Book Company, 1917.

These two books contain not merely well-organized lessons in English but much useful information for foreigners in regard to the geography and industries of this country and the daily life in the city.

G. E. HUBBARD. *From the Gulf to Ararat.* New York: E. P. Dutton and Company 1917. Pp. xv, 273. $3.50.

This is an entertaining account of the activities of the Russo-British commission to establish the Turko-Persian boundary just before the outbreak of the war. As the author travelled over much of the ground since made famous by the disaster at Kut-el-Amara and the subsequent capture of Bagdad, his description throws much light on those gruelling campaigns. The abundance of photographic illustrations with which the book is supplied adds greatly to its charm.

HENRY J. HUMPHSTONE. *Some Aspects of the Memory Span Test: A Study in Associability.* Philadelphia: The Psychological Clinic Press, 1917. Pp. 31.

This monograph is number seven of the University of Pennsylvania Experimental Studies in Psychology and Pedagogy. The purpose of the study was to see how the memory span could be employed for diagnostic purposes. Results were obtained from almost 8,000 boys and girls in six public schools. The author presents detailed tables of his findings, and indicates the distribution of the memory span at each age. The median varied from four at age five to eight for adults. The Pearson coefficient of correlation between memory span and age was 0.4956. There were no sex differences, and no correlation between memory span and school standing.

SETH K. HUMPHREY. *Mankind: Racial Values and the Racial Prospect.* New York: Charles Scribner's Sons, 1917. Pp. xvi, 223. $1.50.

This is a fascinating study of the leading races of western Europe engaged in the present war, and of the fusion of these and other races now going on in America. The author begins with an account of the principles and significance of inheritance, basing his discussion on the work of Galton, Bateson, Pearson, Davenport, Castle, Conklin, and others; considers the essential elements of racial values, the criteria of superior and inferior races; passes to an analysis of the racial traits of the nations now at war; and finally evaluates the racial factors that enter into our own heritage. His comments upon the racial characteristics of France, England, and Germany are trenchant, illuminating, and remarkably free from bias. Especially interesting is his account of the German naivete, consciousness of youthful power, egotism, and insufferable domination. For our own "Melting Pot" the author has only the gloomiest misgivings. Far from developing a new and improved type of civilization, our race mixture can only result in "mongrelism" and decay. Particularly to be

deplored is the race mixture of black and white, which is going on so rapidly that there will soon be no pure blacks left. but a continually increasing infusion of black blood and black traits in the entire population of the land. While not all at flattering to American optimism and complacency, the author's observations are shrewd and biologically well-founded. It is a book to be read and pondered.

JOSEPH JASTROW. *The Psychology of Conviction.* Boston: Houghton Mifflin Company, 1918. Pp. xix, 387. $2.50.

"The subject of this volume is concerned with the interaction of our logicai and our psychological nature. It attempts to deal with the psychology of our most complex logical products. It follows the 'case' method as the only pragmatic procedure, the only one that does justice to the rich content of a concrete issue. In the course of the analysis principles emerge and are emphasized; as in a trial at court, the judge and jury, though concerned with evidence and argument, are guided by principles." The first two chapters deal with the general psychology of conviction, the evolution of belief, the nature, sources and types of credulity, and outline some of the chief cases to be considered in detail. Among these are the belief in the supernatural, with especial reference to the history of Paladino; the history of the study of temperament, with particular attention to phrenology; fact and fable in animal psychology, with consideration of 'malicious animal magnetism' and its incorporation into Christian Science doctrine; the psychology of indulgence, with attention to the crusade against the use of alcohol and tobacco; the feminist movement: and the present tremendous conflict between militarism as represented by Germany and pacifism as set forth by President Wilson. In connection with the latter topic, speaking of the toleration which in normal times is good-naturedly shown to even wild vagaries of belief, the author says: "A flagrant violation of this tolerance appears in the suit instigated by the anti-vivisectionists against the Red Cross organization to prevent the use of funds in the interests of medical research; and that means to mitigate the sufferings and save the lives of the victims of war. To push a private prejudice against a public interest at this time and in this manner is an ignorant, obstinate and malicious attack, inhuman and unpatriotic even though sincere; it is a tragic demonstration of the menace that lies in unreason. Though exceptional, the instance should be used to strengthen the forces of reason and loyalty."

THOMAS JESSE JONES. *Negro Education. A Study of the Private and Higher Schools for Colored People in the United States.* Washington: Bureau of Education, Bulletin, 1916, Nos. 38 and 39. Two volumes. Vol. I, Pp. xiv, 423. $1.00; Vol. II, Pp. v, 724. $1.25.

These two stately volumes with their hundreds of illustrations afford an admirable survey of colored education in the United States. One of the most interesting chapters is devoted to the history of negro education. The second volume describes the facilities for negro education by states.

CHARLES HUBBARD JUDD. *Psychology: General Introduction.* Second edition, completely revised. Boston: Ginn and Company, 1917. Pp. xix, 358. $1.80.

This standard and widely used textbook shows a well-marked evolution in the revised edition. The consideration of conscious activities has been further developed, and more stress is laid on the applications of psychology. Unfortunately the author has not yet emancipated himself from the traditional atomistic conception

of instinct, which allows him to speak of the "development of habit through conflict of instincts." The chapter on speech, one of the best in the book, remains practically unchanged, while that on voluntary action has been much amplified and extended. While the tone of the discussion is strongly behavioristic, consciousness is by no means ignored or belittled.

I. L. KANDEL. *Federal Aid for Vocational Education.* New York: Carnegie Foundation for the Advancement of Teaching, Bulletin No. 10, 1917. Pp. vi, 127.

The first part of this work gives the history of federal aid to education in this country, Part II a detailed account of the constitutional points involved, and Part III recent movements for federal support of agricultural and industrial education.

CALVIN NOYES KENDALL AND FLORENCE ELIZABETH STRYKER. *History in the Elementary School.* Boston: Houghton Mifflin Company, 1918. Pp. viii, 135. $0.75.

It is indisputable that until recently the subject of history has received insufficient attention in American elementary schools. There is no subject which lends itself better to the development of patriotism and a keen appreciation of the advantages and needs of the time in which we live. With the appearance of vastly improved texts in elementary history, and the stress of the present war, the subject should occupy a greater proportionate place in the elementary course. The present little book contains excellent suggestions for the teacher of elementary history. These cover the use of the text-book, the assignment of lessons, study, outside reading, outlines, illustrative material, dramatization, debates, and correlation with other subjects. The book contains many hints for supplementary reading and source materials.

EDWIN A. KIRKPATRICK. *Fundamentals of Child Study.* New edition, revised. New York: The Macmillan Company, 1917. Pp. xxiii, 380.

The principal changes in the third edition of this widely-used book are the transposition and expansion of the chapter on heredity, the omission of the one on the classification of instincts, and the inclusion of selected references from the rapidly growing literature on the subject. The book still remains a treatise on instincts, and in view of the growing skepticism with which psychologists regard this concept it is bound to offer a challenge to the critically-minded. Why the classification of primitive types of behavior was omitted is by no means clear, for certainly no other part of the book has been so widely quoted.

SIR OLIVER LODGE. *The War and After.* New York: George H. Doran Company, 1918. Pp. xiv, 252. $1.50.

"In this book the author takes stock of the past and present and scans the future with relations to the war and to Germany." In the past he notes the increasing materialism and revolt against Christianity in Germany, pictures the prevailing philosophy of life held by the thinkers of Germany, and contrasts fundamental attitudes on life problems in Germany and England. In the present there is a discussion of the nature and effects of war, and the good that can be derived from war if the proper standpoint is taken. In the future the author foresees social and industrial unrest and calls upon the English upper classes to face the issue fairly and through emphasis on a broad education to develop a sense of social solidarity that will conserve the spiritual values of mankind.

CHARLES T. LORAM. *The Education of the South African Native.* New York: Longmans, Green and Company, 1917. Pp. xx, 340. $2.00.

This is the most thorough, scholarly and scientific study of the education of the South African that has yet been made. There is a history of racial relationships in South Africa, an account of the attitude of the whites toward the question of education, a description of the present systems of education, a comparison of the attainments of Europeans, Indians and Natives on the basis of tests in arithmetic, and a plan for a reconstruction of native education. There is frequent reference to the literature of the American Negro, and the progress recently made in Negro education in the South is adequately considered. It is a significant contribution to an important social and racial problem.

G. W. A. LUCKEY. *Essentials of Child Study.* Chicago: The University Publishing Company, 1917. Pp. 219.

This is a series of lecture notes that have grown out of a course in child study that the author has given for the past twenty years. Among the topics considered are the nature and importance of child study, the physical development of the infant, sense discrimination, the feelings, knowing, willing, children's drawings, the beginnings of language, crying and laughing, children's interests, and methods of child study. Valuable features of the book are the extensive lists of references at the end of each chapter and the report of material collected from classes.

EMMA O. LUNDBERG. *A Social Study of Mental Defectives in New Castle County, Delaware.* Washington: Children's Bureau, Publication No. 24, 1917. Pp. 38.

The author finds that of the 212 cases of mental defect studied 82.5 per cent. needed special care, and that over half of these are now at large in the community without any care. The evils of the situation are cogently pointed out, and constructive suggestions are made for improvement.

FRANK E. LUTZ. *Fieldbook of Insects, with Special Reference to Those of Northeastern United States, Aiming to Answer Common Questions.* New York: G. P. Putnam's Sons, 1918. Pp. x, 509. $2.50.

This splendid fieldbook is of proper size to fit the pocket, is conveniently arranged for both scientific and practical needs, and is richly illustrated with plates in color. No more useful or stimulating present could be made to the child just developing an interest in moths and butterflies than this volume. Such a book might well lay the foundation for an illustrious career of scientific investigation.

R. R. LUTZ. *Wage Earning and Education.* Cleveland: The Survey Committee of the Cleveland Foundation, 1916. Pp. 208. Fifty cents.

In this volume we find a summary of the detailed reports on several of the leading industries of Cleveland. The first hundred pages discuss various types of educational organization that may be utilized in the preparation for wage earning. Amongst these are industrial training in the elementary school, the junior high school, technical high schools, continuation and technical night schools, and vocational schools for girls. The volume is an important contribution to both the theoretical and practical aspects of industrial education.

ELMER A. LYMAN AND ALBERTUS DARNELL. *Elementary Algebra.* Cincinnati: The American Book Company, 1917. Pp. vii, 503.

This text does not aim at novelty, but is designed rather to meet the needs of ordinary teachers who expect to drill on the conventional topics, and get as many

of their pupils "through" the college entrance examinations as possible. There is a chapter on graphs, which the teacher is invited to omit if he so desires. Toward the end of the text there is also a chapter on the graphical solution of equations, but it is the exceptional teacher with exceptional pupils who would ever get that far along in the book. The problems are abundant and furnish excellent opportunities for drill.

PATRICK MACGILL. *The Brown Brethren.* New York: George H. Doran Company, 1917. Pp. 296. $1.35.

This is a tale of the trenches, in which the fortunes of a group of London Irish are followed through a year and a half, and the life of the ordinary Tommy is revealed in story form. In the course of the tale the author brings out the varying aspects of modern military life, the humor and good comradeship under privation and suffering, the strain of being under fire, the wild dare-deviltry of the charge, the nonchalant familiarity with death, the relaxation in the billets behind the lines, and the romance with the peasant girls. The book makes the life "over there" very realistic.

F. SCHUYLER MATHEWS. *Fieldbook of American Trees and Shrubs.* New York: G. P. Putnam's Sons, 1915. Pp. xvii, 465. $2.50.

Most educated Americans are singularly ignorant of the life immediately about them. The nature study movement in the schools is making an effort to counteract this by teaching the children some of the more obvious things about their environment, but this is not carried on either very extensively or very intensively. It is a distinct satisfaction to be able to identify the trees that one observes, and this fieldbook will be of great assistance in that effort. There are sixteen full page illustrations in color and fifty in crayon, not to mention the great number of drawings of leaves, fruit, bark, etc. After a little practice one should be able to identify with the aid of this book almost any tree that one comes upon.

MARK A. MAY. *The Mechanism of Controlled Association.* Archives of Psychology, No. 39, August, 1917. Pp. iv, 74. $0.75.

What is the exact mechanism of the control of ideas? This is an experimental study of the nature and effects of preparatory "set." The author distinguishes a fore-period, from the appearance of the task word to the appearance of the stimulus word, a main-period, from the stimulus to the response, and an after-period, continuing for some seconds after the response. The greater part of the investigation was devoted to the time relations of these three periods.

CARDINAL MERCIER. *The Origins of Contemporary Psychology.* New York: P. J. Kenedy and Sons, 1918. Pp. xii, 351. $2.25.

The first edition of this book was issued in 1897, and the present version was translated from the second Belgian edition of 1908. The author first expounds the philosophy and psychology of Descartes, then considers Spencer, Fouillee, and Wundt as masters of contemporary psychology, traces the growth of the experimental movement to 1908, and shows that this has nothing to do with the soul but is rather a branch of anthropology. For the genuine doctrine of the soul Aristotle and St. Thomas Aquinas remain the best authorities, and there is little difficulty in reconciling orthodox belief with the results of experimental investigation. Hence Catholics are urged to participate in these experimental studies, and to enlarge the boundaries of the science of man, rather than to act in opposition to this tendency.

HOLMES W. MERTON. *How to Choose the Right Vocation.* New York: Funk and Wagnalls Company, 1917. Pp. x, 302. $1.50.

The author styles himself "Vocational Counselor," and pretends to suggest "a large number of interesting mental 'tests' which enable one to self-chart one's vocational aptitude." As an illustration, he states, "Form ability makes one conscious of the shape, outline and individuality of objects," and then proceeds to ask a list of questions, such as "Do I actually see the specific projections as well as the general contours and outlines of objects," and expects the reader to decide from his own answers whether he is fitted to be a sculptor, a penman, a photographer or a blacksmith. It is a sad commentary on the intelligence of the public that such travesties have any readers.

C. N. MILLARD. *A Parent's Job.* Boston: The Pilgrim Press, 1917. Pp. ix, 227. $1.00.

Too much of the responsibility for the education of children is shifted from the parents to the schools, according to the author. Therefore it is necessary that the school plan deliberately to enlist the co-operation of the parents, and to induce them to assume their share of the burden. The last half of the book is devoted to a detailed discussion of the school work that should be carried on in grades one to eight.

WALTER S. MONROE. *Second and Third Annual Reports of the Bureau of Educational Measurements and Standards.* Kansas State Normal School: Studies, by the Bureau of Educational Measurements and Standards No. 6, 1917. Pp. 80.

This monograph gives the results of recent studies in arithmetic, spelling, silent reading, visual vocabulary, handwriting, and algebra carried out under the direction of the Kansas Bureau of Educational Measurements.

E. L. C. MORSE. *Spanish American Life.* Chicago and New York: Scott, Foresman and Company, 1917. Pp. 369.

The incidents in this Spanish reader are derived from the best journalistic sources in Spanish-American countries. The selections are taken from Mexico, Cuba, Argentina, and the western coast of South America. They thus give a better idea of Spanish-American life than the ordinary selections from Spanish literature.

E. H. MULLAN. *Mental Status of Rural School Children.* Washington: Reprinted from the United States Health Reports, November 17, 1916. Pp. 30.

This report deals with the results of Binet and other tests in New Castle County, Delaware, and Porter County, Indiana. It is found that in the country, as in the city, a considerable percentage of the children (perhaps two per cent.) are so retarded mentally as to be unable to profit by the regular instruction.

One Year of Compulsory School Attendance in Maryland, 1916–17. Baltimore: State Board of Education, 1917. Pp. 55.

A detailed statistical account of the measures taken for the enforcement of the compulsory education law, and of the results by counties.

A. N. PALMER. *Standards for the Evaluation of Efficiency in Palmer Method Handwriting.* New York: The A. N. Palmer Company, 1917.

This is a series of eight standards to be used one for each grade of the elementary school in estimating proficiency in handwriting. Each sheet contains five or six

specimens ranging from very good to very poor work for that grade, and each sample is evaluated on the basis of posture, movement, speed and formation. The ratings given seem to be entirely arbitary, and probably stand for the best judgment of the author. It is an interesting attempt on the part of a commercial system of handwriting to supply a scale to measure the value of the product.

FRANKLIN C. PASCHAL. *The Witmer Cylinder Test.* Hershey, Penna.: The Hershey Press, 1918. Pp. 54.

This test was adapted as a measure of intelligence from the Montessori cylindrical insets, and the present monograph gives an account of its standardization on 2230 individuals. The method of giving the test is carefully worked out, and extended tables and graphs are given showing the distribution of results. The author considers the test peculiarly adapted to the determination of the degree of distributive attention of the individual.

HENRY CARR PEARSON AND MARY FREDERICKA KIRCHWEY. *Essentials of English.* New York: The American Book Company. First Book, Pp. xii, 308. Second Book, Pp. xii, 454.

The first book is intended for use in the fourth, fifth and sixth grades, and is based on the principle that the best way to train pupils in the use of English is to have them use it. Hence, abundant provision is made for oral and written composition, for practicing the correct use of forms, for letter writing and for the use of different parts of speech. In the second book the idea of applied grammar is carried further, and exercises are given for developing facility in those forms and constructions in which mistakes are frequently made.

BENJAMIN FLOYD PITTENGER. *The Efficiency of College Students as Conditioned by Age at Entrance and Size of High School.* Sixteenth Yearbook of the National Society for the Study of Education, Part II. Bloomington, Ill. Public School Publishing Company, 1917. Pp. 112. $0.75.

High school graduates entering college before 18 do better work than those entering later, and those entering after 19 do poorer work than those entering earlier. Graduates of public schools did better than graduates of private schools. Schools enrolling from 21 to 30 pupils per teacher produced better college students than those with fewer or more pupils per teacher. Women entrants are superior to the men in both scholarship and retention, but there is little difference between those who finish the college course.

MABEL POWERS. *Stories the Iroquois Tell Their Children.* New York: The American Book Company, 1917. Pp. 216.

These wonder and fairy stories were collected by the author from the scattered descendants of the once powerful Iroquois, and are here retold for the benefit of children and grown-ups who are interested in genuine American fairy tales,

Practice Teaching for Teachers in Secondary Schools. Washington: Bureau of Education, Bulletin, 1917, No. 29. Pp. 82. Ten cents.

This monograph is largely the result of the work of a committee of the Society of College Teachers of Education, and was prepared under the supervision of Professor A. R. Mead, chairman of that committee. It contains reports from the leading institutions of the country on the work in practice teaching and studies of the results of this training on the efficiency of teaching and on the pupils taught.

SOPHY H. POWELL. *The Children's Library, A Dynamic Factor in Education.* New York: The H. H. Wilson Company, 1917. Pp. xiv, 460. $1.75.

We are only beginning to appreciate the significance of reading as an educational instrument. Especially in the lower grades the custom is still prevalent of keeping the children for an entire term on a single book, subordinating real interest in reading to drill on the pronunciation of words. Some time we shall appreciate the necessity of having a library in connection with every school, and each child will be expected to read from five to twenty books in a term. The present book will afford valuable information for the selection of these libraries. There are chapters on libraries for elementary schools, high schools, rural schools, the relation between the public library and the schools, the training of the children's librarian, and the selection of books for children. A very useful part of the book is the bibliography, extending over more than 100 closely printed pages.

H. O. RUGG. *The Cost of Public Education in Grand Rapids.* Pp. 361–475.

This is an extract from an extended report on the Grand Rapids schools, and gives tables and graphs of expenditures for education in Grand Rapids compared with those of other cities of like size. It is valuable for students of educational finance.

LEROY W. SACKETT. *Comparable Measures of Composition.* Reprinted from School and Society, Vol. 5, 1917, 233–239.

The author gives an account of grading a series of compositions by four different methods, and contends that better results are achieved by a score card than by either the Hillegas or the Ballou Scales.

ANNA TOLMAN SMITH AND W. S. JESIEN. *Higher Technical Education in Foreign Countries. Standards and Scope.* Washington: Bureau of Education, Bulletin, 1917, No. 11, Pp. 121. Twenty cents.

An account of the organization and equipment of the higher technical schools of the chief countries of Europe, Asia, and South America. A very important document in view of the industrial reorganization that may be expected after the war.

ALICE TEMPLE. *Survey of the Kindergartens of Richmond, Indiana.* Supplementary Educational Monographs, No. 6. Chicago: University of Chicago Press, 1917. Pp. v, 58. $0.40.

This survey brings out once more the absurd fashion in which many communities handicap their bright children. By a state law no children in Indiana are allowed to enter the first grade until they are six years old. But many bright children entering the kindergarten at four or four and one half are ready for the first grade in one year. Instead of advancing according to their capacity, these children are compelled to mark time, with all of its deadening monotony. Of course the enterprising superintendent might give these children first grade work in the kindergarten and advance them correspondingly in the elementary grades, but this is not easy to manage in crowded schools. The surveyer found the customary lack of articulation between the kindergarten and the first grade, chiefly because of the inflexibility and unnaturalness of first grade teaching.

LEWIS M. TERMAN. *The Intelligence Quotient of Francis Galton in Childhood.* Reprinted from the American Journal of Psychology, 28: 1917, 209–215.

On the basis of the data presented in Karl Pearson's *Life, Letters and Labors of Galton,* Professor Terman finds that Galton was distinctly an infant prodigy, with an intelligence quotient of perhaps 200.

VOLUME IX OCTOBER, 1918 NUMBER 8

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

AN ANALYTIC SCALE OF HANDWRITING

C. C. LISTER AND G. C. MYERS
Brooklyn Training School for Teachers

In 1910 a new epoch in educational measurement began, when Thorndike presented the first scale for measuring school products.* It was a writing scale. This scale was so carefully constructed that it has served as a model for practically all the scales appearing since that time. He had one thousand specimens "ranging from the best to the worst handwriting found in grades 5 to 8" arranged by from twenty-three to fifty-five competent judges, each grading the set into groups 1 to 11 by what he considered equal steps.

Two years later Leonard P. Ayres† constructed a second scale. This scale was made from 1578 samples of the handwriting of children of the upper elementary grades of forty school systems in thirty-eight states. Ten paid investigators made accurately timed readings of each sample. The legibility was considered the measure of quality.

These two very carefully constructed scales have been widely used for measuring handwriting in the public schools. In spite of their many merits they have certain defects, and do not wholly meet the needs of the schools. Some of these defects have been pointed out by Freeman.**

The chief criticism from the viewpoint of the practical school teacher is that while these scales are more precise in measuring writing product than hit-or-miss judgment without a scale, neither

Handwriting. Teachers College Record, March, 1910.

†*A Scale for Measuring the Quality of Handwriting of School Children.* 16 pp., 1912. Division of Education, Russell Sage Foundation.

***An Analytic Scale for Handwriting.* Elementary School Journal. 15, 1915: 432–441.

helps the teacher to analyze the child's writing defects. For Ayres, "general quality" is the only guide, and the assumption is that the general quality is measured by speed of reading, or legibility. Thorndike aims to judge three factors combined, legibility, beauty and character. After offering a number of criticisms on his own work he says (p. 13–14): "A far more sagacious criticism than either of these would be that a scale like this for merit in general is less useful than a scale for legibility alone, or for beauty alone or for character alone, or for ease alone. Of course, I admit that such specialized scales are highly desirable, and I hope that this scale for general merit will stimulate others to the labor of making similar scales for legibility alone, beauty alone, and so on. But it seems sure that the scale of most importance and usefulness is that for general merit."

Now suppose one had such specialized scales and were to rate Willie's writing thereby so as to find it 8 in beauty, 16 in legibility and 12 in ease; one would have to say, "Willie, your writing is quite legible but you must make it more beautiful." From the viewpoint of any of the leading systems of teaching penmanship this diagnosis would be of little value to the teacher or the child. On the other hand there are certain specific essentials which are looked for by leading penmen. Freeman suggests that the chief ones are uniformity of slant, uniformity of alignment, quality of line, letter formation and spacing. Accordingly he set out to construct a scale from these five viewpoints.

In making his scale Freeman constructed a preliminary chart presumably of artificial specimens, to illustrate different degrees in these traits. "This was used as a guide by a class in experimental education composed of advanced students, most of whom were experienced in supervision and teaching, upon which to rank a large number of specimens into ten degrees of excellence in each trait. After the specimens had been so rated, however, it appears to the writer that the order was not always the correct one." Therefore Freeman proceeded to "doctor" the results, and as objective measures of slant and of alignment "it was only necessary to measure the angles of a series of letters and to find the mean variation among these angles, or to measure the vertical positions of the tops and bottoms of the letters and to measure the variability among those positions. When this was done it was found that the order based on variability as measured did not correspond to the order on the basis of the judgment made by the grades. The order based on

objective measurement was therefore used as a basis for the selection of specimens for the scale." "In the case of quality of line no means was found upon which to base such objective measurement; but the characteristic in question was made more prominent by photographic enlargement. When this was done it was relatively easy to determine differences in the irregularity of the line of the writing. On this basis specimens for this chart were selected."

In the case of letter formation the assistance of a Mr. S. "who had developed a system of determining excellence in letter formation by the method of counting the errors in form" was called to note "the errors in the specimens of the scale which had been selected. In some cases the results of this measurement differed from the results of the judgment of the graders, and a compromise between the two methods of determining was used in selecting the specimens for the scale."

Of spacing he says: "Various specimens were constructed in which the spacing between letters was different, and the judges asked to select that specimen in which the letters were the most agreeably spaced" . . . . "Then the lower grades were constructed by varying this spacing in a variety of ways." For the scale, specimens of writing that compared most closely with these standards were chosen.

"Each chart contains specimens of writing which represents three grades of excellence in the characteristics in question." The lowest is valued 1; middle, 3; highest, 5; and double weight given to each rank in letter formation.

Certainly Freeman set out to meet a definite need; but from the description of his mode of procedure there is evidence of considerable looseness. For example, "class in experimental education," "a large number of specimens," "a series of letters," "various specimens," and "varying this spacing in a variety of ways" are not terms suggestive of scientific accuracy. Apparently the preliminary chart was artificial. So far as judges were concerned this preliminary chart was really the scale. There is no evidence to show how this was constructed.

For slant and alignment, the opinions of the judges were arbitrarily thrown out of court and instead "a series of letters" one knows not how many or whether letters of children or of adult were actually mechanically measured in respect to regularity of slant and

alignment. This type of measure is perhaps in itself not without merit; but the shifting of standards is of questionable propriety.

For this as in succeeding traits Freeman notes that "On this basis specimens for the chart were selected." Just who selected the final specimens is not always clear. In the case of letter formation it seems that a Mr. S. was the final arbiter. Finally no statement is made as to how the three specimens used on the scale were selected from the "ten degrees of excellence" rated by the "class."

Assuming that Freeman's scale were sufficiently scientific, it does not quite meet the needs which Freeman so clearly has pointed out. Three degrees of merit can hardly suffice in diagnosing writing with great precision, and in grading as accurately as most teachers are called upon to grade.

Furthermore it appears that no one has attempted to construct a scale from a product of any one recognized school system of penmanship. Indeed, Thorndike criticizes his own scale because it does not represent all styles of writing. Ayres, however, has arranged his specimens on the scale in three rows illustrating slant, medium and vertical writing. Apparently the specimens, as presented to his "readers," were all mixed and those on the scale merely selected from the general results. A few words from Ballou in his study of Boston's spelling are to the point.

"In conclusion, the writer urges the fundamental importance in the standardization of any educational product, or in the evaluation of questions in a standard test for any subject, of knowing that the pupils have received instruction in the field of knowledge covered by the scale on the test. Unless scales on the standards derived from standard tests are based on the results achieved by children after proper instruction, then standards established on the basis of the results from such tests are unsatisfactory as measures of instruction. That Boston pupils spell much better than pupils in the 84 cities is evident from evidence here submitted. To the writer the reason for this superiority appears to be the fact that Boston pupils have received instruction in the words under consideration; whereas, the standards in the Ayres' scale are based on results obtained from pupils who may never have been taught to spell the words in the scale. To assume that the scale establishes proper standards by which to measure the results of spelling instruc-

tion is to be satisfied with less ability among pupils than may reasonably be expected if pupils have been properly instructed."*

To incorporate the leading merits of the writing scales noted above, and to supply some of their defects, the writers have attempted to construct a scale. In 1914 a uniform style of letters to be taught in New York City was adopted by the Board of Education, and teaching of the muscular movement method of writing in the city schools was authorized. Under these favorable conditions the writers, at the invitation of Acting Superintendent Straubenmuller, proceeded to construct a writing scale for New York City. The following which briefly states the mode of procedure appears at the bottom of this scale:

How the Scale Was Made

The scale represents the average judgment of 21 teachers and penmen expert in the muscular system of penmanship, and 4 psychologists. From 9 schools of Greater New York, representative of the best, medium and poorest product of the muscular system of penmanship 3550 specimens were selected from at least one entire class of each grade from 3B to 8B inclusive. Each specimen represented one trial from dictation.

The specimens from each grade were classified into four piles by the writers on the basis of general merit and each pile was thoroughly mixed with its corresponding pile of the several grades. Then on the basis of chance three hundred specimens were selected so that practically the same number was drawn from each pile. According to written instructions each of the 25 judges ranked these specimens in 8 piles on the basis of equal intervals in merit. Accordingly each judge ranked the 300 specimens three times, namely, as to form, spacing and movement.

On the basis of the average rank assigned each specimen the best and the poorest were selected as the top and the bottom of the scale. Therefrom the exact numerical rank which the other six samples should have was determined. The specimens whose average ranks are the same as these determined positions, or are nearest them were selected. Without exception all the samples on the scale are less than .1 from the *determined* position.

In addition it should be noted that one of the writers (C. C. L.) is personally familiar with the general nature and quality of the writing in practically every elementary public school of the city;

---

*Frank W. Ballou. *Measuring Boston's Spelling Ability by the Ayres Spelling Scale.* School and Society 5: 1917, 720.

that the writing was selected from 3 schools of Manhattan, from 3 of Brooklyn and from one of each of the other three boroughs and that at the outset of selecting the 300 specimens, about a score of papers which were blotted or were too dim for photo-engraving, were discarded.

Here is a sample of the directions given:

INSTRUCTIONS TO JUDGES

These specimens will be handled by about forty judges. Therefore you will please handle them as carefully as possible.

Please make all judgments with perfect independence. Any suggestions from another judge will render the work unreliable.

First. Arrange the 300 specimens in eight piles according to their order of merit in *movement,* so that the intervals between the piles will be as nearly equal as possible. The basis of judgment as to movement will be the quality of the lines. Heavy tremulous lines indicate poor movement. Frequent lifting of the pen between the letters in a word indicates poor movement. Sharp, smooth, clear-cut lines indicate good movement.

As soon as this is done, copy the numbers placed on the specimens, on the blank sheet labeled *Movement.* The numbers found on the papers in the pile of best specimens should be grouped after *One;* those on the papers in the second pile should be grouped after *Two,* etc.

Second. Mix the specimens (take one from each pile and put it on a new pile till all are collected) and then arrange them in eight piles according to their merit in *form.* Form includes accuracy in letter formation, uniformity in size and uniformity in slant.

When the papers have been classified as to *form,* record the numbers on the blank sheet labeled *Form,* as directed above for recording movement.

Third. Again mix the specimens and arrange them in eight piles according to their merit in *spacing.* In judging spacing consider correctness and regularity of distance between letters and words; spacings that are too wide, too close or irregular are undesirable.

When the samples are arranged as directed tabulate the numbers on the sheet *Spacing,* after the manner indicated above.

Be sure to sign your name to each record sheet.

Your co-operation will be appreciated.

Three types of these directions were provided in which the essentials appeared, movement—form—spacing, form—movement—spacing, spacing—form—movement orders respectively. Each direction sheet was used about as often as every other, though unfortunately exact distribution was not recorded. Several who gave introspections noted that judgment from one point of view did not affect judgment from other points of view, yet there was doubtless some effect. Whether it would have been better to have had 3 sets of judges instead of the one no one knows. This, however, should be determined experimentally.

Before each judge began his work these directions were read over with him to make sure he was certain of how to proceed. He was especially urged to strive to make the steps between the piles equal. The judges were told before they volunteered that the task would take from 4 to 8 hours, and the purpose of the study was explained to them. While two reported that they did it in a "sitting" most of the rest noted that they did only one ranking at a time.

The data indicating the determined position which each specimen on the scale should have for movement, form, and spacing,

TABLE I

Form

Sample Rank	Sample Number	Desired Value	Real Value	Probable average divergence of the estimated quality from an estimate by an infinite number of judges.
1	30*	1.20	1.20	.06
2	135*	2.152	2.17	.15
	287		2.12	.12
	150		2.20	.21
3	3*	3.104	3.12	.14
	38		3.17	.14
	82		3.17	.20
4	167*	4.056	4.04	.15
	218		4.04	.21
	11		4.08	.17
	16		4.08	.13
5	109*	5.008	5.00	.11
	94		5.00	.15
	281		5.00	.20
6	225*	5.960	5.96	.14
	160		5.96	.15
	162		5.95	.15
	290		5.95	.15
7	222*	6.912	7.00	.09
	226		7.00	.10
8	53*	7.87	7.87	.04

respectively, are given below; also the samples whose actual rankings equal or approximate the desired rankings. Those starred are the samples selected for the scale. Furthermore, the probable average divergence of the concensus of our 25 judges from the concensus of opinion of thousands of such judges is indicated below, along with similar data from Thorndike.

TABLE II

Movement

Sample Rank	Sample Number	Desired Value	Real Value	Probable average divergence of the estimated quality from an estimate by an infinite number of judges.
1	36*	1.12	1.12	.04
	42		1.12	.04
	30		1.12	.04
2	92*	2.084	2.00	.18
	291		1.96	.15
3	43*	3.048	3.08	.15
	130		3.08	.25
4	111*	4.012	4.00	.14
	122		4.00	.15
	275		4.00	.19
	192		4.04	.15
	211		4.04	.15
	248		4.04	.14
5	294*	4.976	4.96	.13
	109		4.96	.19
	117		5.08	.20
	178		5.08	.15
6	281*	5.940	5.92	.16
	159		5.92	.18
	84		5.92	.24
	289		5.96	.16
	73		5.96	.17
	68		5.96	.18
7	226*	6.904	6.92	.11
	217		6.92	.12
	134		6.92	.11
8	116*	7.87	7.87	.04

TABLE III

Spacing

Sample Rank	Sample Number	Desired Value	Real Value	Probable average divergence of the estimated quality from an estimate by an infinite number of judges.
1	37*	1.25	1.25	.07
2	46*	2.214	2.12	.16
	9		2.32	.15
	176		2.37	.16
3	12*	3.178	3.16	.14
	129		3.16	.16
	107		3.20	.17
	192		3.22	.16
4	20*	4.142	4.13	.15
	194		4.13	.13
	131		4.12	.24
	218		4.12	.25
	23		4.12	.17
	140		4.12	.24
	171		4.12	.17
	286		4.12	.17
5	76*	5.106	5.09	.16
	203		5.12	.18
	155		5.08	.21
	250		5.08	.17
	206		5.04	.14
	273		5.04	.15
	184		5.04	.24
	178		5.04	.24
	259		5.04	.21
6	282*	6.070	6.08	.18
	5		6.08	.18
	295		6.09	.21
7	115*	7.03	7.04	.16
	105		7.00	.12
	268		7.00	.10
8	182*	8.00	8.00	0

TABLE IV

Thorndike's Data (p. 11 Handwriting)

Sample	Quality	Probable average divergence of the estimated quality from an estimate by an indefinite number of judges
32	16.1	.14
84	16.2	.43
47	15.0	.19
49	15.1	.18
89	15.0	.39
90	15.1	.35
19	14.0	.20
54	14.0	.19
4	12.9	.20
24	13.1	.18
26	12.9	.18
55	13.1	.21
30	11.9	.19
7	12.0	.20
52	12.0	.20
23	11.0	.20
45	11.0	.19
106	11.0	.28
17	10.2	.18
21	9.1	.15
28	8.9	.15
31	8.9	.14
48	8.0	.14
14	8.1	.19
126	7.0	.40

As it happened, several specimens received the same average rank. In that event the specimen upon which the judges most closely agreed (indicated by the probable average divergence) was selected. There is one exception to this rule: for position number four under *Spacing*, the specimen from the pair equally nearest the true position, which was selected had the larger probable divergence. The other, however, had been slightly crumpled from handling, whereby it was rendered poor for photo-engraving. When, as in a few cases, the probable divergence as well as the average position was the same for several specimens (e.g. numbers 1 and 7 of *Movement*), photo-engraving merit determined the choice of the specimen for the scale.

It is rather remarkable that so many specimens should come to have the same rank. Indeed, those specimens with practically any one of the numbers listed under each of the eight positions could have been taken as scale samples without serious departure from the truth. The fact that the same system of writing was employed by those writing the specimens may suggest why so many samples were ranked on the average, the same. Mere inspection of the product of the muscular system of penmanship indicates that individuals write much more like one another than they would write, probably, were the system used not uniform.

The closeness of the selected specimens to the determined positions (less than .1) is noted in the data printed on the scale. On this point Thorndike says (p. 10) "As was noted, on page 3, the scale is only approximate. 16 on the scale does not pretend to mean 16.00000, but between 15.9 and 16.1. 8 does not pretend to mean 8.0, but between 7.9 and 8.1, and as a matter of fact, although I have had a thousand samples graded and have chosen as wisely as I could, some of the samples do vary in merit from 7, 8, 9, 10, etc., by more than .1 plus or minus." (See data from Thorndike below.) Thorndike also points out that a new but small varying element is introduced by the process of photo-engraving and printing.

Although these data are not easily comparable with those of Thorndike since he used more degrees in ranking, had more specimens and more judges, they are considerably lower and suggest as great, if not greater, reliability than do his. There are certain reasons why for the same number of judges our data should admit of slightly greater reliability: the judges looked for specific things; nearly all were expert penmen in the style of writing judged; the

FORM

Form includes accuracy of letter formation, uniformity of slant and size.

Beautiful penmanship,
like elocution or music is an ac-
complishment that naturally at-
tracts all people of refinement

The above model shows the style of writing taught and illustrates ideal letter formation, size and slant.

90

One must exercise in work and in
play Active play and almost all kin
of work which children have to do, a
good forms of exercise. Long walks
once or twice a week are good, but

This is good elementary school penmanship. Note the uniformity of size, slant, and alinement. The x is crossed carelessly.

80

One must exercise in work and in
play. Active play and almost all kind
of work which children have to do are
good forms of exercise. Long walks
once or twice a week are good, but

This writing is too small; but note the uniformity of slant and size. It is too angular—see m, n, and h.

70

One must exercise in
work and play Active
play and almost all kinds
of work which children
have to do, are good forms

Good form. The curvature between letters is slightly exaggerated. 'Note the approach to a, d, and m.

60

One must exercise in work and in play. Active play and almost all kinds of work which children have to do are

This writing is too angular, especially m and n. It slants too much. Note the careless form of k.

50

One must exercise in work and in play Active play and almost all kinds of work which children have to do, are good forms of exercise. Long

This writing slants too much. The t should not be looped. Note the careless tendency in completing k.

40

One must exercise in work and in play. Ative play and almost all kinds of work which child-ren have to do, are good forms of exercise. Long walks once

The slant is irregular. The r's look like i's and the s's are round at the top. The x is crossed carelessly. The f is poor.

30

One must exercise in worke and in play. Active play and almost all kinds of work which children ha to do. are good forms of

This writing slants too much. The loops are too long and narrow. The t should not be looped. The a and d are poor.

20

One must exercise on work and

Active pleay and, almost all king of

wrchod. childre ndo, good fost loow

once are twice a week, but their are n

gut, so good as regular exercise d

specimens were all in one style of writing and the same subject matter was involved. On the other hand Thorndike had more judges than we had and they judged 1000 instead of 300 specimens. Our specimens, however, were chosen by lot from 3550.

More judges would have reduced the probable error but the specimens would have been endangered by wear and slight crumpling. Future makers of such scales should have specimens written on cardboard or on paper mounted in some way.

At the top of the scale under movement are the standard letters as adopted by New York City, and, at the top of form and spacing columns respectively are model specimens from a copy book, arbitrarily placed there to illustrate the ideal toward which the teacher and school should work.

It will be noted that the writers have paid no attention to the zero point on the scale. Thorndike strongly emphasized this and Ayres also. Likewise scale makers in other subjects have worried over it. The attempt by them of course, has been to make a scale analogous to physical measures. This attempt doubtless was of rare value in developing confidence in scales when the first scale was born; but for practical purposes the writers agree with H. T. Manuel,* that "elaborate effort to establish a zero point is quite unnecessary," and that "any two samples may be taken as fixed points."

We have attempted merely to show that the specimens selected as the eighth and first specimens on our scale are those selected as the highest and lowest from our specimens, and nothing more.

According to the probable average divergence, the indications are that the steps for the several essentials between the first and second position, and between the seventh and eighth position, are greater than the other intervals on the scale. Especially does this seem true with the eighth specimen of spacing, which was ranked in that position by every judge. Indeed a number of the judges stated that this specimen for spacing "spoiled the equal steps." Had the specimens to be judged been selected from a much larger number this difficulty might have been slightly decreased. In the light of the findings of a number of investigators using the ranking method, however, irregularity of intervals is practically inevitable, especially for the extremes. For practical purposes

*_The Use of an Objective Scale for Grading Handwriting._ Elementary School Journal, January, 1915.

this fact is of little importance in respect to the function of the scale. When, however, arbitrary relative values are assigned to the steps on the scale, as the writers have done, the error is slightly exaggerated. It will be noted that the sample for the lowest position is assigned a value indicating it to be more than a "step" below number 7. Of the shortcomings of these arbitrary values the writers are aware, but these values are merely offered as suggestions to the many teachers who insist upon having some such guides. Certainly, wherever the administrator permits, all grading of the child's writing would better be done in terms of mere position on the scale.

On the scale appear simple concrete directions for its use. Under each specimen are a few suggestive remarks analyzing that specimen. Freeman gives grades 1, 3 and 5 respectively to each of the three points on his scale for each essential except in letter formation, where double weight is assigned each rank. To arrive at relative weights one could determine as above, a general merit scale and thereby ascertain to what degree movement, form and spacing respectively correlate with this general quality.

With the same three hundred specimens which are used in this scale, but properly mounted for protection, the writers hope later to make such a study. The results thereby obtained would not wholly suffice as relative weights for all grades alike, since the writing product per se is hardly the equally desirable objective in all the grades. From the writer's viewpoint, at least, the first few grades should emphasize the ground work of muscular movement penmanship and therefore put relatively greatest emphasis upon movement. One obvious difficulty with this obtains; the child as well as the adult likes to get practical results. If, however, the proper standards are set for the child and teacher any desirable objective could be made to become "practical." In the absence of sufficient data the writers would recommend further that equal weights be assigned the three essentials in the fifth and sixth grades and that more weight be given to form and spacing and less to movement in the seventh and eighth years. This scale is of doubtful value below the 4th grade.

To determine how the writing of the various grades is distributed on our scale is another bit of unfinished business.

STANDARD GEOGRAPHY TEST—THE WORLD
For Fifth Grades

ERNEST C. WITHAM

Southington, Conn.

The geography test here described was worked out two years ago, and since then has been given in several parts of the country. Standard is used in connection with this series of tests to distinguish them from the old line of examinations. The author regrets exceedingly that his administrative duties have prevented his carrying out some plans of standardization which he has had in mind regarding the development of his geography tests. Many requests for samples and information regarding the tests is the only excuse for offering this article at this time.

Purpose of the Tests.

The general purpose of these tests is to enable those responsible for the quality of instruction to measure the work done by the pupils and teachers in the subject of geography. These tests offer a reliable means of getting a large body of facts, and a record of geographical thinking in a minimum amount of time. They are also intended for wide-awake individual teachers, who are looking for school room helps.

Directions for Giving "The World" Tests.

For supervisory purposes it is best to have the same person give this test in all the different fifth grade rooms; but this is not absolutely necessary. Where several are to give the test, they should all be given general instructions beforehand, so as to make the conditions the same.

This test should be given in the early part of the spring term to fifth grades. By this time the pupils should know the geography called for. If it turns out that the class as a whole is not able to make a good score on the test, it is not too late in the year to begin to remedy the defects. Individuals who are below standard should be given special attention. This test will be especially helpful in diagnosing their difficulties. Every pupil should be given a set of the test papers. A few general words of advise should be given to the pupils in regard to carefulness, neatness, and honesty. The pupils should be told not to hurry as they will be allowed all the time they need to complete the tests. (The usual time is about 25 minutes for all to complete the fifth grade test.) The pupils should next fill in name, school, and room at the top of the page.

THE WORLD—FOR FIFTH GRADES

I. On the outline maps printed below find the following geographical divisions of the world, and write the name of each across the face of the map.

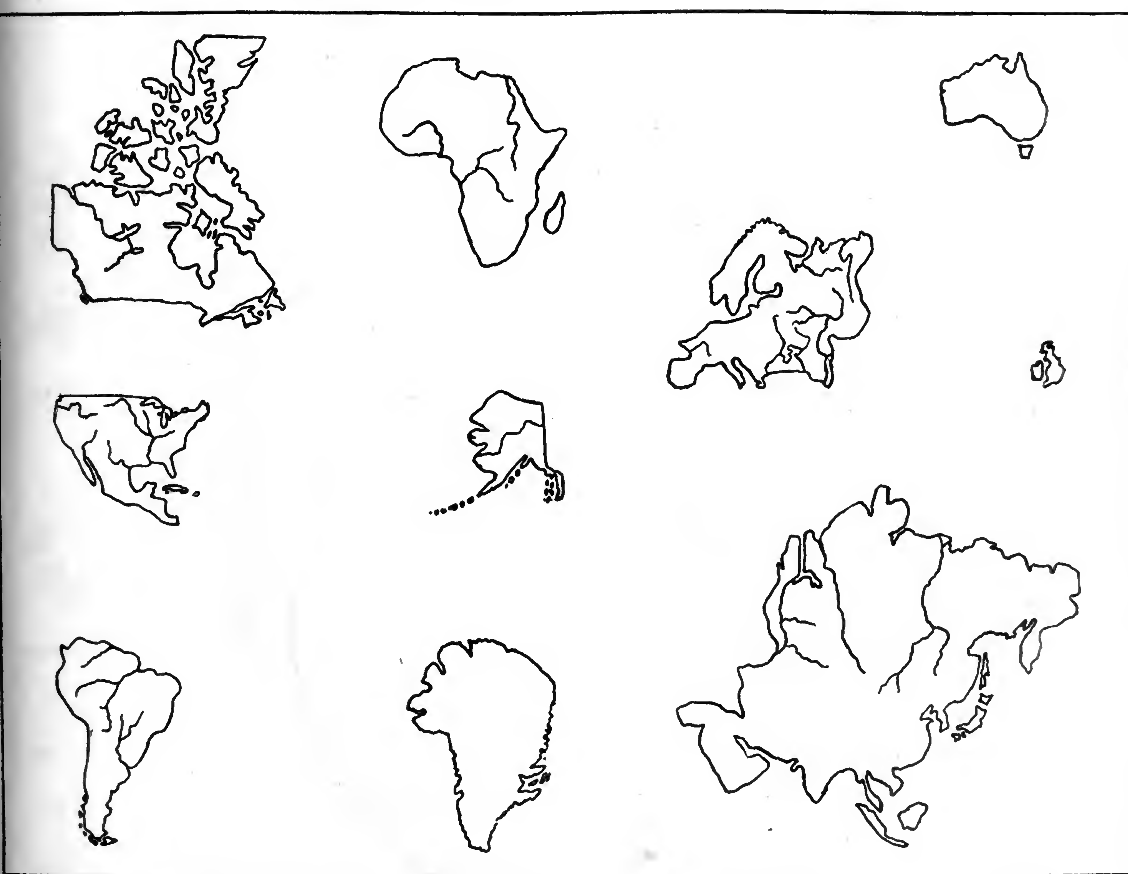
1. The United States, Mexico and Central America.
2. Greenland.
3. Dominion of Canada.
4. South America.
5. Alaska.
6. Great Britain (British Isles).
7. Africa.
8. Asia
9. Australia.
10. Europe.

II.

1. What is the form of the earth?....................................
2. What are the motions of the earth?..............................

III. On the maps below:

1. Draw the Equator.
2. Indicate the North Pole.
3. Indicate the South Pole.
4. Write under each map which hemisphere it is.
5. Write in their proper place the names of the following oceans:
   a. Atlantic.
   b. Pacific.
   c. Indian.

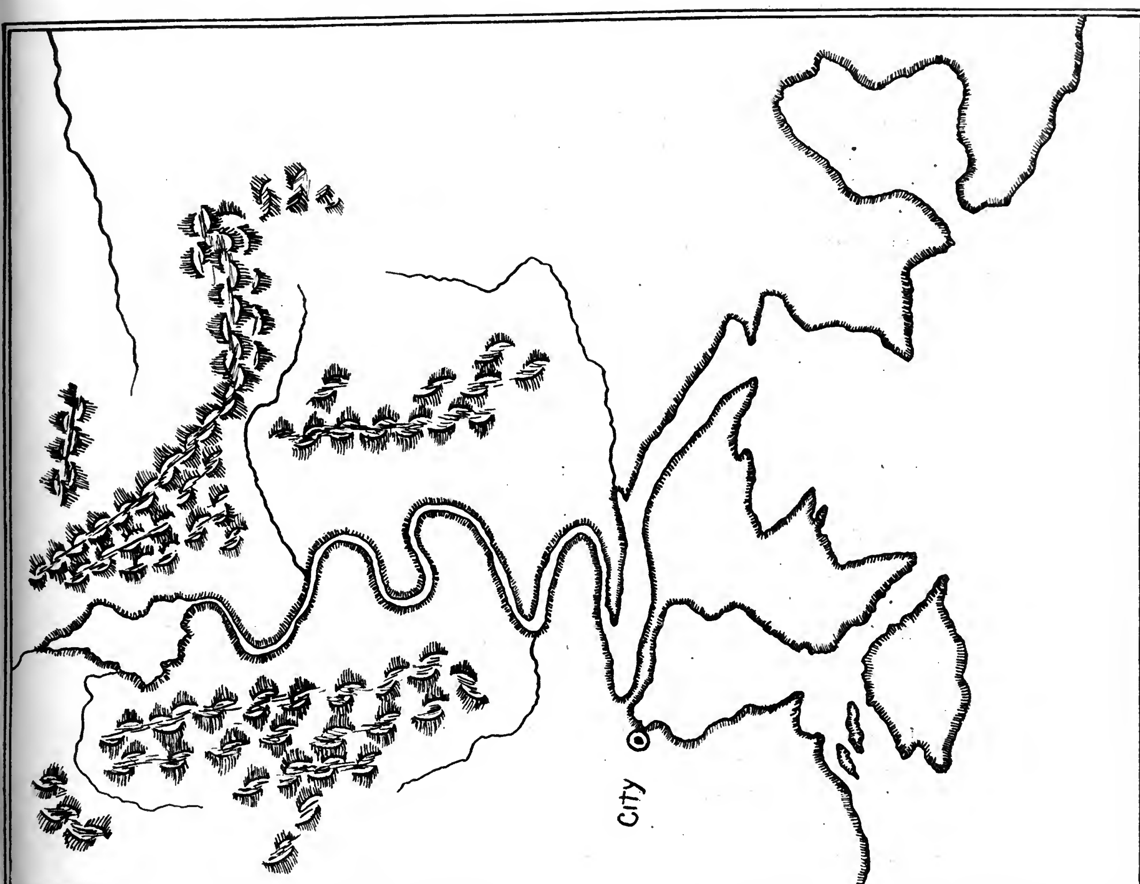


----------Hemisphere.

----------Hemisphere

IV. Indicate on this map all of the following land and water forms:

Land Forms	Water Forms
1. Island.	8. Lake.
2. Mountains.	9. Gulf.
3. Peninsula.	10. Bay.
4. Cape.	11. River.
5. Valley.	12. Ocean.
6. Mainland.	13. Strait.
7. Isthmus.	14. Harbor.



V. What continents are the homes of the following peoples?

1. Red..
2. White..
3. Yellow...
4. Black..
5. Brown..

VI. Name one great industry of each of these countries:
1. United States..
2. Argentina..
3. Canada...
4. France...
5. Russia...

VII. The five largest cities in the world are given in order of their size. In what countries are they located?
1. London...
2. New York...
3. Paris..
4. Tokio..
5. Chicago..

All pupils should now read aloud the first question. The examiner should illustrate just what is wanted. This can be done best by quickly outlining a rough map on the blackboard. Any one of the states will do. For example, Maine. The examiner will then say, "Now this is not one of the geographical divisions called for, but if it were, you would write the word Maine, like this, right across the map. Now turning to your paper, you will notice ten geographical divisions called for. Study them over carefully, and then write the name of each across the map, just as I have done in the case of the map of Maine, which I have roughly represented on the board."

As fast as the pupils complete the first page have them tear it off and pass it in. Pupils should go ahead with the other questions as fast as they finish the first. They must be reminded to write their names on each page before beginning to answer the question. They should be told that there are two parts to the second question under section II. In section III the pupils should be told to write the names of the Atlantic and Pacific oceans on both maps.

Directions for Scoring "The World" Test.

First sort the papers having all of the first pages in one pile, the second pages in another pile, etc. Correct the first pile of papers, and mark in the rectangle in the upper right hand corner the number of answers attempted, and also the number of answers correct. Next correct the second pile of papers. In Section II, answers such as—sphere, like an orange or like a ball, are correct. Just the word, round, should be called wrong. There are three possible points in this section.

Section III.—Ten possible points.
Section IV.—Fourteen possible points.
Section V.—Eight points.

Answers are as follows: (The order of the answers does not matter.)
Red 1. North America 2. South America

White 3. Europe 4. North America 5. South America
6. Yellow—Asia—Also Japan and China is a correct answer.
7. Black —Africa.
8. Brown—Asia—also call East Indies or Philippines and Hawaiian Islands correct answers.

Section VI.—Five points.

Call correct any of the following answers:
1. United States—Farming, manufacturing, lumbering, mining, cotton, fruit growing, cattle raising, etc.
2. Argentina—Cattle raising, wheat.
3. Canada—Lumber, wheat, furs.
4. France—Manufacturing, iron, linen, silk, porcelain, glassware.
5. Russia—Grain, or wheat, cattle, lumber.

After the papers have all been corrected, take the first pile, and sort the papers into groups according to the number of answers attempted. For example, if there are 30 papers, place all with ten attempted in one pile, all with nine attempted in the next pile and all with 8 attempted in another pile, and so on. Record the number attempted on the Class Record Sheet. F means frequency. If there are 20 papers with ten attempts in section I, write 20 in column F, opposite 10, which is the score (Sc). If there are 5 papers with nine attempts, write 5 opposite the 9. If there are 2 with seven attempts mark a 2 opposite the 7. If there are 3 with six attempts write 3 opposite the 6. The second column is for the sum of the scores. In each case multiply the frequency by the score. Add the first and second column and record at the bottom. Divide the sum of the second column by the sum of the first. This will give the average number of points attempted in section I.

Assemble the first pages and sort again according to the number of correct answers. Count the number in the several piles and record on the right side of Section I of the Class Record Sheet, column F. Proceed as before to get the sum and average. The remainder of the tabulations should be carried out in the same manner as in the case of Section I.

Graph Sheet. To get a class graph make an X as near as possible at the point on each section of the graph sheet which indicates the number of attempts. Connect these with a solid line. Similarly, construct a graph showing the number of rights, using a broken line.

A few of the interesting replies to the test questions follow:

Fighting was given by several pupils as a great industry in both France and Russia.

"Industries of Argentina are fruit and earthquakes."

One pupil located Paris in Germany and another located Chicago in Michigan.

Only a few instances showed lack of ability to read the questions. The whole test was arranged with this point in view.

Figures 1 and 2 show two of the class records, giving the distribution of results of the papers which were scored according to the directions.

I

F	S	Sc	F	S
33	330	10	20	200
		9		
2	16	8	3	24
		7	3	21
		6	2	12
		5		
		4		
		3	2	6
		2	4	8
		1	1	1
		0		
Sum 35	346		35	272
Ave.	9.9			7.8

II

F	S	Sc	F	S
3	9	3		
30	60	2	13	26
1		1	20	20
1		0	2	
Sum 35	70		35	46
Ave.	2.			1.3

III

F	S	Sc	F	S
31	248	8	4	32
2	14	7	3	21
1	6	6	10	60
		5	3	15
1	4	4	6	24
		3	1	18
		2	1	2
		1	1	1
		0		
Sum 35	272		35	173
Ave.	7.8			4.9

IV

F	S	Sc	F	S
12	168	14		
3	39	13		
6	72	12	6	72
4	44	11	1	11
5	50	10	3	30
2	18	9	4	36
1	8	8	5	40
1	7	7	3	21
1	6	6	3	18
		5	3	15
		4	4	16
		3	1	3
		2	2	4
		1		
		0		
Sum 35	412		35	266
Ave.	11.8			7.6

V

F	S	Sc	F	S
32	160	5	5	25
3	12	4	14	56
		3	8	24
		2	6	12
		1	2	2
		0		
Sum 35	172		35	119
Ave.	4.9			3.4

VI

F	S	Sc	F	S
25	125	5	5	25
4	16	4	4	16
3	9	3	13	49
3	6	2	5	10
		1	6	6
		0	2	
Sum 35	156		35	106
Ave.	4.5			3.

VII

F	S	Sc	F	S
31	155	5	3	15
2	8	4	8	32
		3	14	42
		2	1	2
		1	6	6
2		0	3	
Sum 35	163		35	97
Ave.	4.7			2.8

FIGURE 1

Explanations for Figures 1 and 2

In the tables in Figures 1 and 2 the central column (Sc) indicates the score, the columns marked F indicate the frequency or number of pupils making that score, and those marked S show the totals, or the frequency times the score. The totals on the left give the number attempted, those on the right the number correct.

I

	F	S	Sc	F	S
	15	150	10		
	7	63	9		
			8	1	8
	1	7	7	1	7
	3	18	6	1	6
	3	15	5		
	1	4	4		
			3		
			2	6	12
			1	12	12
			0	9	
Sum	30	257		30	45
Ave.		8.6			1.5

II

	F	S	Sc	F	S
			3		
	18	36	2	12	24
	8	8	1	13	13
	4		0	5	
Sum	30	44		30	37
Ave.		1.5			1.2

III

	F	S	Sc	F	S
	19	152	8		
			7		
	2	12	6	2	12
	4	20	5	6	30
	1	4	4	6	24
			3	11	33
	3	6	2		
			1	1	1
	1		0	4	
Sum	30	194		30	100
Ave.		6.5			3.3

IV

	F	S	Sc	F	S
	6	84	14	1	14
	3	39	13		
	2	24	12	2	24
	2	22	11		
	2	20	10	1	10
	2	18	9		
	1	8	8	1	8
	2	14	7	1	7
	3	18	6	3	18
			5	3	15
	2	8	4		
	2	6	3	6	18
			2	6	12
			1	3	3
	3		0	3	
Sum	30	261		30	129
Ave.		8.7			4.3

V

	F	S	Sc	F	S
	18	90	5		
	3	12	4		
	4	12	3		
	4	8	2		
	1	1	1	17	17
			0	13	
Sum	30	123		30	17
Ave.		4.1			.6

VI

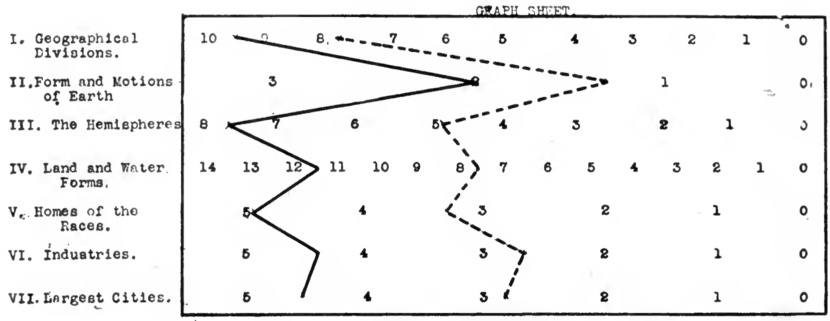
	F	S	Sc	F	S
	4	20	5		
	4	16	4	1	4
	5	15	3	2	6
			2	4	8
	2	2	1	3	3
	15		0	20	
Sum	30	53		30	21
Ave.		1.8			.7

VII

	F	S	Sc	F	S
	6	30	5	1	5
	3	12	4	2	8
			3	2	6
	3	6	2	1	2
	2	2	1	4	4
	16		0	20	
Sum	30	50		30	25
Ave.		1.7			.8

FIGURE 2

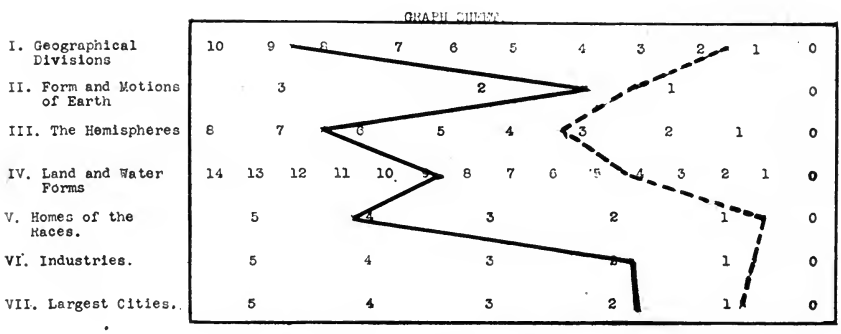
Figures 3, 4, and 5 are graphs of the results of the test in three fifth grade rooms. All the tests were given and scored by the same person. Figure 3 is the graph of the results as given in figure 1, and figure 4 is the graph of the results as given in figure 2.



Draw a solid line to show the number of attempts, and

a broken line to show the number of rights.

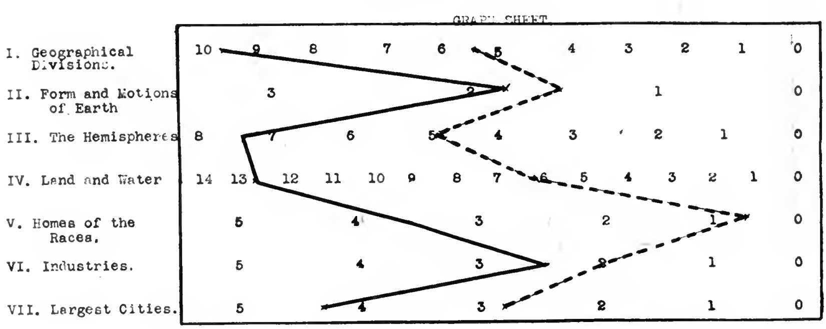
FIGURE 3



Draw a solid line to show the number of attempts, and

a broken line to show the number of rights.

FIGURE 4



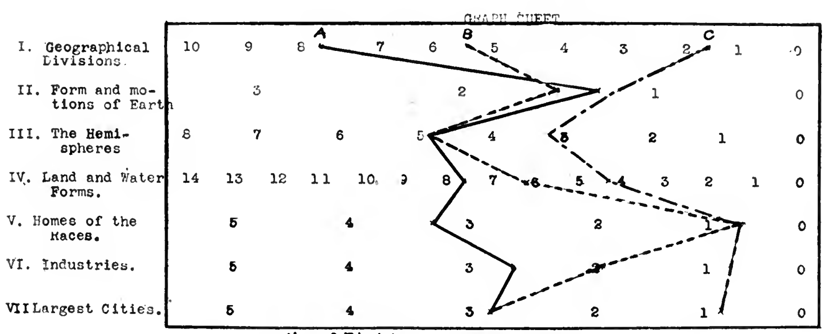
Draw a solid line to show the number of attempts, and

a broken line to show the number of rights.

FIGURE 5

There can be very little question as to which of these classes was getting the best teaching in geography. It is also equally clear which class was next best. Since the use of these tests I can safely say that the class where the poorest work was being done is no longer at the foot of the list. The test proved the right stimulus at the right time and the result has been good. This particular teacher has had years of experience, but believed that she must push the pupils through every page of the geography regardless of everything else. The test has helped wonderfully in bringing about a better method of teaching geography.

Figure 6 shows the scores of three pupils in the school that made the best record. The number of right answers of the best, the median and the poorest pupils in the class are graphically shown.



No. of Rights in Three Schools.

FIGURE 6

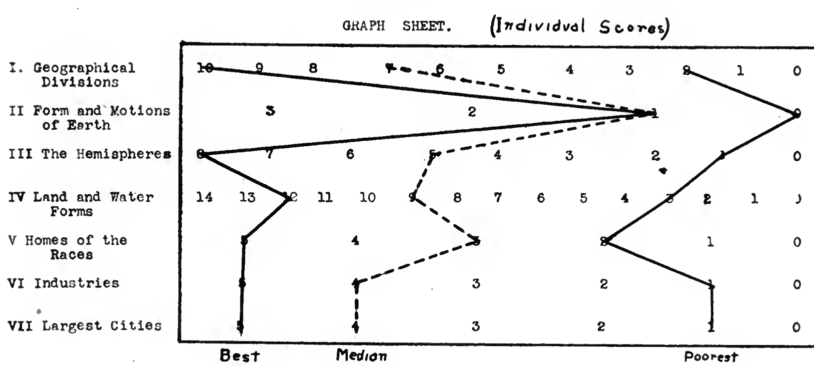


FIGURE 7

The pupils in the room showing the best results were arranged according to their rank in the standard test, and their relative positions are shown by the numbers in the last column of figure 8 In the first column of this same figure the pupils are arranged according to their rank on the results of test No. 20 "The World"

Thompson's *Minimum Essentials.* In the second column the pupils are ranked according to the teacher's marks in geography. So that beginning at the top of the figure the first pupil ranked 34 in the Minimum Essential test, 3 in the teacher's marks and 15 in the standard geography test. The second pupil came 6 in the Essential test, 30 in teacher's marks and 20 in the standard test, etc. The connecting lines show graphically the relationship between these three sets of ratings.

The Pearson coefficients of correlation for the results in the Essential test, teacher's marks and standard test are as follows:

Standard test and teacher's marks $= +.41$
Standard test and Minimum Essentials $= +.58$
Teacher's marks and Minimum Essentials $= +.57$

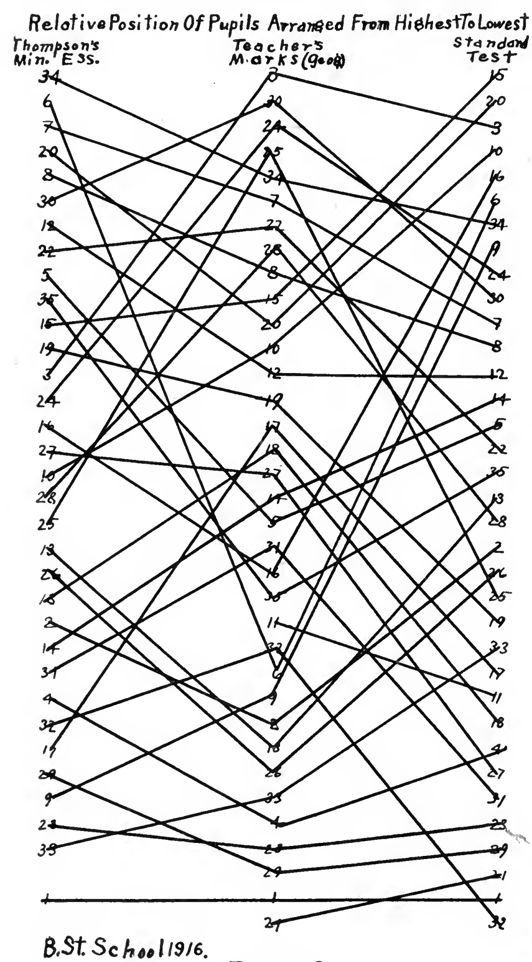


FIGURE 8

A SCALE FOR MEASURING THE ABILITY OF CHILDREN IN GEOGRAPHY

E. E. LACKEY

State Normal School, Wayne, Nebraska

The scale herein described is a product of the joint efforts of H. H. Hahn, Dean of the Department of Education, and E. E. Lackey of the Department of Geography, both of the State Normal School at Wayne, Nebraska. The latter assumes responsibility for the selection of questions, for what constitutes acceptable answers and the adequacy of the tests as a whole to fairly cover the field of geography. The science of scale-making and the classification of the exercises needed in the construction of this scale is the contribution of the former. Although each assumes responsibility for the contribution from his distinctive field yet it was found that each could assist the other in almost every phase of the work. The following description of the derivation of the scale is the work of Dean Hahn.

Purpose of Scale

Since texts will be used by a large majority of teachers for years to come, our primary purpose was to construct a scale for the testing of the teaching of geography from text-books. But when we realized that not one but a number of texts are being taught, we had to modify our plan. Our first modification consisted of limiting our exercises to the phases of geography treated in common by six modern texts. Then we found that some of these phases were treated more fully by some authors than they were by others. A second modification of our plan was, therefore, necessary; namely, to select the common subject matter, or, in other words, the essentials of subject matter in each phase. In the selection of the essentials of subject matter the common subject matter in these texts was largely our guide; but we also checked our exercises by principles and minimum essentials as they have been worked out by makers of geography curricula. (See 1914 and 1916 Year Books of the National Society for the Study of Education.) Over six hundred questions and exercises were selected by three teachers, covering this common subject matter. These questions and exercises were then examined by the authors of the scale, first with reference to repetitions,

and duplications were eliminated. They were next examined for language difficulty. The wording of many of the exercises was changed, some of them were actually tried out on children, and in many cases technical expressions which would convey exact meaning to mature students of geography were eliminated and the ordinary language of children substituted. This is particularly true of the exercises in the lower reaches of the scale. The exercises intended for the upper reaches of the scale were not freed from technical expressions the meaning of which pupils are expected to know as evidence of geography ability. Thus we find such expressions in some of the exercises of the scale as "the Fall Line," "climate," "continent," "natural wonders," "natural geographic barriers," "agencies," "cyclonic storms," and many others equally technical. The exercises were examined, in the third place, as to their scope, as suggested before. Nothing was included beyond the essentials of geography. Finally, the list of exercises was revised so that it contained about an equal number of memory and thought questions and exercises.

The scale, therefore, does not test the teaching of any one text. It only tests the teaching of that subject matter which it has in common with the other five texts. But while it is not a complete test for the work done in any one text, it has the compensatory virtue of being a test, more or less, of the minimum essentials of geography so far as they are determined under the above limitations.

The Preliminary Test

The exercises were mimeographed in sets of twenty-five each, with a place for the answer after each exercise. They were given to pupils in the fourth, fifth, sixth, seventh, and eighth grades. Their instructions were as follows:

1. Write the answer to each exercise in the spaces directly following it. Some exercises call for two or more things. Be sure to give as many things as the exercise calls for. Write merely the answer. Do not use complete sentences.

2. If you cannot answer a question at all, leave spaces for answer blank. If you cannot answer because you do not know what the exercise means, write "Do not know what it means."

3. Ask no questions about any of the exercises in the test. Should you forget and ask questions, your teacher must refuse to answer them. If your teacher should permit you to ask questions and then

answer them for you, it would defeat the purpose of the whole test and your answers could not be used. So be sure not to ask a single question, but do the best you can in answering each question as it is given in the test.

4. Work as fast as you can, but do not get in a hurry. Be sure you know what each exercise means before you begin to answer it.

5. At the bottom of each sheet write your name, grade, school, town, date, and the time the whole class begins a test and the time when *you* finish it.

Scoring of Answer-Papers in Preliminary Test

Pupils were not limited as to time. The test was given to 1696 pupils in twelve schools in two states. The answers, 283,000 in number, were all graded by the authors of the scale. They worked out the correct answers to the exercises together, and then scored the papers accordingly. Where an exercise consisted of two or more parts, credit was given for each part answered correctly. Credit was also given for answers that were somewhat incomplete clearly on account of language difficulty. Each set of questions was scored on the number of exercises answered correctly, number answered incorrectly, number not answered, and the number not answered because the meaning was not understood. The percent of correct answers to each exercise in each of the five grades for each of the twelve schools was determined and tabulated. (These tables are ready for inspection or publication, but are not included in this article.)

Derivation of the Scales

Scale A

In deriving scale A the school grades were ignored. The collective judgment of all of the pupils—1696 of them—independent of school grade and school training, was taken on each exercise. On the basis of this collective judgment, the exercises were ranked so as to represent different grades of geography difficulty ranging all the way from the maximum amount of difficulty expressed on the scale by "0" to almost no difficulty expressed by "100." Since the collective judgment of 1696 individuals is probably not able to designate more than twenty-five grades between the minimum and maximum amounts of geography difficulty, the exercises were arranged into twenty-five groups or steps. In order to get the differences between these steps to represent equal differences of geography difficulty

or geography ability, it was necessary on the assumption that geography ability is distributed among school children in harmony with the law of chance, to make the percentile values of the twenty-five groups of exercises equal respectively to the values of a like number of equal divisions of the surface of the so-called normal probability distribution. Accordingly the base line of the normal distribution surface was divided into twenty-five equal parts and their values to the nearest whole number determined in terms of the surface divisions. These twenty-five values are the following: 0, 1, 2, 4, 6, 8, 12, 16, 21, 27, 34, 42, 50, 58, 66, 73, 79, 84, 92, 94, 96, 98, 99, 100. These values were taken, then, as the values of the twenty-five groups of exercises, and each exercise on the basis of the rank given it by the collective judgment of the children was placed in its proper step of the scale. The exercises that did not have any of these values were not used. Some of the exercises had only approximately the values of the steps into which they were respectively placed. To be exact, in no case does the absolute value of an exercise deviate from the approximate value by more than four-tenths of a step, and this only in a very few cases. Expressed in per cent., this means that 50 represents values from 46.8 to 53.2; 34 represents values from 31.2 to 37.2; 27 represents values from 24.6 to 29.8; and so on through the other values of the scale. In the construction of scale A, exercises have been considered as representing equal geography difficulty if the collective judgments they received by the pupils in all the grades were equal, an assumption that is probably not true and that will not be necessary in deriving scale B. If our assumption be true, then all the exercises in each step of scale A are equal in geography difficulty. The differences between consecutive steps of the scale are also approximately equal, made so by dividing the base line of the surface of normal distribution into equal divisions, as explained above. We have, then, in scale A, a valid means of measuring ability of children in geography. However, the weakness of this scale lies in using for its construction the collective judgment of children in different school grades and therefore of unequal amounts of geography training. This possible error is eliminated in both scales B and C.

Scale B.

The method used in the construction of scale B is a modification of the one used by Dr. E. L. Thorndike in his derivation of *A Scale for Handwriting of Children in Grades 5–8.* A brief description of

this method is found in Strayer and Thorndike's Educational Administration, pp. 208–226. The modifications are these: first, in place of using "from 23 to 55 competent judges" to rank the exercises in order of geographical difficulty, we used from 18 to 21 seventh and eighth grade classes representing 659 different pupils; second, instead of depending upon the judgment of "competent judges" for the equality of exercises placed in the same group, as to geography difficulty, we used the actual performance of school children under normal conditions; and, third, instead of trusting the opinion of "competent judges" to get the geography difficulty between successive groups equal, we used the twenty-five values obtained by dividing the base line of the normal distribution surface into twenty-five equal parts, as was described in connection with the derivation of scale A.

In the construction of scale B the exercises were ranked into twenty-five groups in order of geography difficulty, the same as in scale A. The values of the twenty-five groups or steps are the same as in scale A and were determined by an identical process. The rank given each exercise by each class was determined by the average performance of the members of the class. The final rank given each exercise in the scale by all the classes combined was determined by their median rank. For instance, exercise 185 was placed by the twenty classes that judged it on the basis of average performance, in group 12 once, in group 13 two times, in group 14 three times, in group 15 eight times, in group 16 three times, in group 17 two times, and in group 18 once. The median rank of the twenty classes is group 15, and in group 15 exercise 185 was placed in the scale. The place of each exercise in scale B was determined in this way.

It may be argued that since seventh and eighth grade classes were used in placing these exercises in scale B the same error is involved in its construction as in scale A; namely, the error due to unequal amounts of geography training. So far as actual performance is a criterion the seventh and eighth grade classes reveal, in the preliminary test, almost identical geography ability. Furthermore, nearly all of the schools that were tested teach very little geography in the eighth grade. The test was given to these schools about a week before they closed the year's work. The seventh grade classes had practically finished their course in geography and had,

therefore, approximately the same amount of training as the eighth grade classes. The little additional training the eighth grade classes received was offset by the more recent training in the seventh grade. Thus the fact that the pupils used in the construction of scale B belonged to two different school grades did little or no harm. But even this objection, harmless as it is, is eliminated in the construction of scale C. (Neither scale A nor scale B has been published, but a copy giving merely the number of the exercises in each step will be furnished upon request.)

Scale C

Scale C is the result of a combination of Thorndike's method, modified as described in connection with scale B, and Ayres' method of scale construction. It is not a single scale, but a multiple scale, combining scales for grades 4, 5, 6, 7, and 8. The groups or steps and their values are identical with those in scales A and B, and were determined by identical methods. To rank the exercises in order of geography difficulty in each single grade scale, only pupils of that grade were used, thus eliminating the objection of unequal amounts of geography training. The principle is generally accepted that exercises can be said to be equal in difficulty only when they are equal in difficulty for pupils who have had equal amounts of training.

The multiple scale C was derived as follows: first, single scales were constructed for grades 4, 5, 6, 7, and 8; secondly, the five single scales were then combined to form the multiple scale C for the five grades. The single scales were constructed in the same way as scale B with the exception of using pupils of one grade only. In combining the five single scales the seventh grade scale was used as a basis for comparison. To determine the relative position of the single scales in combination it was thought best to find the average deviation, in terms of scale-steps, of the ranks of the different exercises as they appeared in each of the single scales for grades 4, 5, 6, and 8, from their corresponding ranks in the single scale for grade 7, the one used as the basis for comparison. For instance, in comparing the position of exercise 220 in the scale for grade 6 with its position in the scale for grade 7, it was found that there is a deviation of two steps; that is, exercise 220 is placed in the former scale two steps nearer the "0" group than in the latter scale; exercise 221 deviates two steps, exercise 222, three steps; exercise 223, two steps, and so on, the total deviation of all the exercises

being four hundred twenty-five steps, or an average deviation of approximately two steps. The average deviation between the ranks of the exercises in the scales for grades 5 and 7 was found to be approximately three steps; between the scales for grades 4 and 7, approximately five steps, and between the scales for grades 8 and 7, one-tenth of a step. The average improvement from grade to grade, reduced to terms of scale-steps, shows exactly the same step differences as does the average deviation. Using, then, the seventh-grade scale as the basis for combining the five scales, we placed the scale for the eighth grade immediately below that for the seventh grade and made their steps to coincide; the scale for the sixth grade we placed above the seventh and two steps to the right, next the fifth-grade scale and three steps to the right of the seventh; and last the fourth-grade scale and five steps to the right of the seventh. This, then, constitutes the multiple scale C, and is the one published under the title, *A Scale for Measuring Ability of Children in Geography in Grades 4, 5, 6, 7, and 8.*"

Correlation of Scales A, B, and C

The correlation between scales B and C is .99 with an average variation of less than .5 of a step. This shows that the two scales are almost identical. Either one of these two scales is exact enough for measuring ability of children in geography. Scale A shows an average variation of approximately two steps from scales B and C, thus proving that the objection mentioned in its description is valid.

Representative Parts of Scale C

The following is a part of scale C, "knocked down" and re-arranged so it can be presented in straight pages as a part of this article. The May, 1918, standards for the various grades are given in connection with each step. To illustrate: "Step S—58% (4); 73% (5); 79% (6); 88% (7); 88% (8)" means that in a test on the exercises in Step S the fourth grade should average 58%; the fifth grade, 73%; the sixth grade, 79%; the seventh grade, 88%; and the eighth grade, 88%. Steps G and H are comparatively difficult exercises, Step M is of average difficulty and Steps R and S are comparatively easy. The exercises in Roman type are the ones which test the memory and the ones in italics are thought-provoking.

Step G—1% (4); 4% (5); 6% (6); 12% (7); 12% (8).

207. Name three agencies or processes at work making rocks into soil.

215. By what states would you pass in going by boat from Cincinnati to Memphis?

150. *Why is the rainfall of Australia limited to the eastern and southeastern parts?*
162. *Much of India receives from 12 to 16 inches of rainfall in July and less than 1 inch in January. Explain.*
225. *Which is the greater distance and why, 30° west of Washington or 30° south of Washington?*
216. *New Orleans is in 30° North Latitude and St. Louis is in 39° North Latitude. They are in the same Longitude. About how far apart are they in miles?*

Step H—2% (4); 6% (5); 8% (6); 16% (7); 16% (8).

123. Which way does your shadow point or extend just before sundown in mid-summer? In mid-winter?
152. How does Russia in Europe compare in size with the United States?
198. Give four marked illustrations of man's skill in overcoming natural geographic barriers.
221. Between what two bodies of land is the Bering Sea? Dover? Skager Rak? Babel Mandeb?
110. *What is the "Fall Line," and why are a number of cities located on it?*
144. *Why is there a heavy rainfall in the Amazon valley?*
157. *What is the cause of rain?*
177. *The heat equator passes through New Mexico in July and through Argentina in January. Give the chief reason why?*
200. *Give the principal reason why the coastal cities of the United States have a more temperate climate than cities of the same latitude in the Great Central Plains.*
191. *It is noon at Omaha. State the time at the following places: Baltimore, Denver, New Orleans, San Francisco.*

Step M—16% (4); 27% (5); 34% (6); 50% (7); 50% (8).

90. How could you go to Asia if you wished to make that trip?
93. Name two large rivers of Asia.
111. How can you go by boat from the Hudson River to Lake Erie?
133. Give two reasons for the importance of the Columbia River.
136. Name two of the most important materials shipped on the Great Lakes.
155. Give capitals of Japan and China.
159. Name five important inland cities of Europe.
212. Draw a map of your own state and locate in it two rivers, the capital, and the largest city.
214. Name the state or territory in which each of the following is located: Galveston, Washington, St. Paul, Sitka, Savannah, Spokane.
199. *What disadvantage do the people of Great Britain suffer as to food supply?*
109. *Why do the rivers of New England furnish water power for manufacturing?*
135. *Why is the Rio Grande an important river?*
138. *Give two reasons why cities usually grow up at waterfalls.*
140. *Give two reasons why Argentina exports wheat to Brazil rather than to the United States.*
165. *What part of Asia is similar to Canada and in what way is it similar?*
188. *Since the larger part of our iron ore is mined in Minnesota, why is little iron and steel manufactured there?*
117. *Why does the earth not look round to us?*

76. *Why are so many hogs raised in the United States?*

Step R—50% (4); 66% (5); 73% (6); 84% (7); 84% (8).

62. Name the five Great Lakes of North America.
73. What are the two largest cities of the United States?
29. Name two ways in which the farmer helps to get food for us.
41. Give two ways in which water gets away when it rains.
3. What is the name of the circle extending around the earth midway between the poles?
98. *Give one reason why the people in the far north use reindeer and dogs instead of horses.*
37. *Why does not the water flow out of swampy places?*
46. *Why do not the Eskimos build houses like ours?*
6. *What is under the ocean?*
19. *Where does the water in a well come from?*
23. *Name two ways in which winds are useful.*

Step S—58% (4); 73% (5); 79% (6); 88% (7); 88% (8).

52. What is the largest city of your state?
641. Where is Alaska and to whom does it belong?
84. Name four large cities of Europe.
92. Give the capitals of France and Germany.
101. Name two large bodies of water that border on Florida.
45. Name four things you use for food that do not grow where you live.
68. *Give one reason why so many of the great cities of the United States are near the sea coast?*
72. *Which is the coldest and which the warmest part of South America?*

THE VOCABULARY TEST AS A MEASURE OF INTELLIGENCE

LEWIS M. TERMAN

Assisted by S. C. Kohs, Mary B. Chamberlain, Mayme Anderson and Bess Henry
Leland Stanford Jr. University

Criticisms of the Test

Perhaps no mental test in current use as a measure of endowment meets such instant criticism from laymen and theoretical psychologists as the vocabulary test. Scores of times the writer has been very positively informed that vocabulary has nothing to do with intelligence, and that even if the two were correlated the relationship could not be established by a test consisting of only a hundred words. The following are the most common objections to the test:

1. That the number of words known depends upon accident of environment and instruction, not upon intellectual endowment; that to know the significance of a vocabulary score we should have to know the subject's home, the number of years he had attended school, the quality of speech used by his playmates, the number of books read, etc. The average person would regard the use of the vocabulary test with a school child whose parents speak a foreign language as the limit of folly for mental testing, even if the subject had spoken English since early childhood.

2. A second criticism of the test is that any innate ability it may measure is a very special ability, not general intelligence. It is sometimes asserted that this special ability is *negatively* correlated with intelligence. Psychologists themselves have often contrasted the "verbal" type of individual with the "logical" type, to the disadvantage of the former. It is a common opinion that there exist feeble-minded persons, of the so-called "fluent" type, who have immense vocabularies.

3. The critic is also sure to insist that even if the vocabulary test were valid in principle, no dependence could be placed on the score obtained from a 100 word list selected at random from a dictionary containing 18000 words.

4. Finally, the test is believed by some to be largely invalidated by the personal equation in scoring.

In view of such criticisms we have taken the trouble to assemble here a few facts bearing on the validity of the test in question. What

has been done is only a beginning, and it is to be hoped that someone will before long make a thorough quantitative and qualitative study of vocabularies.

Correlations with Mental Age

The best way to measure the reliability of an individual test is to correlate it with the total score from a group of standardized tests the reliability of which is known. We have used the Stan-.

TABLE I

Vocabulary Score

Mental Age	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	Total
19													1	1	2	1		5
18^6													5		2	1	1	9
18										1		1	1	2	3	2	1	12
17^6											1	2	6	3	3	2	1	18
17										1	2	2	9	1	1			16
16^6										2	6	3	5	2	2			20
16										6	8	13	10	3	2	1		43
15^6									4	8	9	8	5	1	3			38
15					1			2		16	11	9	4	2				45
14^6								1	4	11	11	5	4			1		37
14								3	6	11	6	1	1	1				29
13^6					1		2	5	6	7	2	1	1					25
13						3	2	7	2	6	5						1	26
12^6					1	2	2	6	4	4	3							22
12				1	1	3	3	10	4	1	2							25
11^6						3	5	6	2	1			1					18
11				1	1	11	10	2	1	1								27
10^6			1	1	6	4	1	3		1								17
10			2	2	6	14	5	1										30
9^6	1	1	1	5	4	11	1											24
9		1	3	11	3	7	2	1										28
8^6		2	10	6	5													23
8	1	6	4	9	2													22
7^6	1	6	4	5			1	1										18
7	4	13	9	2														28
6^6	5	11	2															18
6	1	5																6
5^6		2																2
Total	13	47	36	43	31	58	34	48	33	77	66	45	53	16	18	8	4	631

$r = .91$

Note. Mental age 9–6 equals 9–6 to 9–11, inclusive, etc.; vocabulary score 10 equals 10 to 14, inclusive, etc.

ford Revision of the Binet-Simon Scale for this purpose. It is not claimed that this scale is a perfect measure, but the degree of its reliability is fairly well established. Mr. Otis has devised a method* for calculating the probable error of a Stanford-Binet mental age score. Application of this method in a large number of cases shows the probable error to be less than 6 months in terms of mental age when the test is used with miscellaneous adults, and about 3 months when used with children of the first school grade. When the scale is split endwise and only three tests are used in each year, the probable error is still only about 7½ months. Any single test which correlates closely with such a test series may be regarded as having a high degree of reliability.

Table I shows the correlation between vocabulary score and mental age earned on the Stanford Revision. The 631 subjects were school pupils scattered from grade I to the first year of high school. The intelligence quotients ranged from less than 50 to more than 150. All but a few of the children were from homes where English is spoken.

Very few tests or groups of tests yield as high correlations as that found in the above table ($r = .91$), but it is the usual thing for the vocabulary test. It is evident that a mental age based on vocabulary score alone would not be far wrong in a large per cent. of cases. We have determined the probable error of such a mental age, which we may call the "vocabulary mental age," and found it to be approximately 9½ months. The following table shows the probable error of a "vocabulary mental age" for various vocabulary score ranges:

*Described in The Psychological Clinic. 1918.

TABLE II

Vocabulary Score range	P. E. of mental age for each score range
10–14	.6 year
15–19	.6 "
20–24	.6 "
25–29	.9 "
30–34	.8 "
35–39	.8 "
40–44	.8 "
45–49	.9 "
50–54	.8 "
55–59	.8 "
60–64	.6 "
65–69	.9 "
70–74	1.1 "
75–79	1. "

Average P. E. .8 year = 9.6 months

From the above we can infer that when we are dealing with school children mental age based on the vocabulary test alone will not deviate from a mental age earned by the entire Stanford Revision more than

9½ months in 50 per cent. of cases.
12 " " 40 " " " "
18 " " 20 " " " "
24 " " 10 " " " "
36 " " 1 " " " "

Table III shows the correlation of vocabulary with mental age in the case of 482 miscellaneous adults, including:

150 "Hoboes" tested by Mr. H. E. Knollin
150 Prisoners tested by Mr. H. E. Knollin
150 delinquent youths tested by Dr. J. Harold Williams
32 business men

TABLE III

Vocabulary Score

Mental Age	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	Total
19													1	1	1	1	1	5
18^6													2	4	3	2		11
18												1	5	3	4	3		16
17^6									1			1	2	2	3	1		10
17									1	3	1	6	3	2	4	1	1	22
16^6									1	1	5	3	2		2			14
16							2	2	3	5	2	5	4	1	3			27
15^6						1	1	2	5	4	5	3	5	4	1	1		32
15							3	4	3	6	5	1	1		2			25
14^6							4	5	6	11	2	3	2		1			34
14						3	8	8	7	2	2	1	2	1				34
13^6						7	8	7	5	3	2	2						34
13				3	5	3	7	12	5	3	2	2						42
12^6				2	1	8	7	8	2	1	1							30
12				2	6	6	8	5	1	1	1							30
11^6			1	3	7	5	1	7	1									25
11		1	1	6	8	6	5	1										28
10^6	1		1	5	4	4	3		1									19
10	1		2	9	4	2												18
9^6			2	1			2											5
9	1	3	3	2	1													10
8^6	2	1	2		1													6
8		1	1															2
7^6			1															1
7			1	1														2
Total	5	6	15	34	37	45	59	61	42	40	28	28	29	18	24	9	2	482

$r = .81$

The correlation in the above table is .81 Pearson. This is high, but not quite as high as that found for children. The difference is readily accounted for by the extraordinarily motley character of the group which contained individuals of many races, all degrees of education and quite a number who had spoken another language before learning English. For this group the P. E. of a mental age based on vocabulary approximates 12 months. A mental age thus secured would not, even for such subjects as these, deviate from a Stanford-Binet mental age more than:

12 months in 50 per cent. of cases.
18 " " 31 " " " "
24 " " 18 " " " "
36 " " 4.3 " " " "

Intelligence Quotient and Vocabulary

It has been shown that vocabulary correlates highly with mental age, or the absolute mental level. We have also raised the question whether it bears any constant relation to intelligence quotient, which is an index of *relative* brightness. In order to answer this question we have divided the 640 school children into three groups, including respectively those with I. Q. below 86, 86 to 114, and above 114. The median vocabulary of each of these three groups was as follows for the various mental ages:

TABLE IV

Mental Age	Below 86	86–114	Above 114
17–6 to 18–5			74
16–6 to 17–5		66.9	65
15–6 to 16–5		61.9	56.9
14–6 to 15–5		57.0	55
13–6 to 14–5		51.7	50
12–6 to 13–5	46.7	47.5	43.5
11–6 to 12–5	43.7	41.4	40.2
10–6 to 11–5	34	33.7	36.2
9–6 to 10–5	32.5	29	30.6
8–6 to 9–5	24	22.2	21.7
7–6 to 8–5	18.3	18	20
6–6 to 7–5	13.7	12.3	12.5
No. in Group	112	150	185

It will be seen that at no mental age do the three groups differ considerably in median vocabulary. The only constant tendency noticeable is for the group of highest I. Q. to fall slightly below the middle group after the mental age of 12 years. We have found this tendency quite marked in children above 140 I. Q.

The last mentioned fact deserves emphasis. We have often been told that our subjects who test so unusually high, say 140 or above, are probably less bright than they seem; that they belong to the "verbal" type and test high on the Stanford-Binet because it is so largely a language scale. The reverse of this is actually the case. Such children do less well on the language tests of the scale than they do on tests which make heavier demands upon reasoning. Mr. Kohs has found that children of exceptionally high I. Q. by the Stanford-Binet, consistently earn a still higher I. Q. by his Block Design scale, a scale which is made up entirely of performance tests.

It is also worthy of remark that although the children of a given mental age who are below 86 I. Q. are from one to several years older than those of the median and bright groups, this advantage has practically no effect on the vocabulary score. The latter depends upon mental level and is but little influenced by chronological age.

If age apart from intelligence affected the vocabulary very much the fact would certainly appear in the median vocabulary scores of Knollin's 482 miscellaneous adults. These were mostly from five to forty years older than the "mental age" earned in the Stanford-Binet test, yet as is shown by the figures in Table III their vocabulary scores by mental age were only slightly above those of school children of the same mental age, except at the highest levels, where the discrepancy is somewhat more marked.

TABLE V

Median vocabulary scores of children and adults by mental age

Mental Age..	7	8	9	10	11	12	13	14	15	16	17	18	19
Children..	12.9	18.7	23	31	34	41.9	46	51.5	56.8	61.7	66.1	73	75
Adults..			26	33.5	38.7	44.1	49.6	52.3	61.5	68.5	73.9	83	85

Note. Mental age 7 = 6–6 to 7–5, etc.

The correlations with mental age were computed separately for the three groups, bright, average and dull. They were as follows:

Group	*Correlation (Pearson)*
Above 115	.95
86–115	.944
Below 86	.86

Effect of Foreign Language in the Home

In the last two or three years some three hundred Portuguese and Italian children have been tested by various Stanford students. Most of these children were from homes in which Portuguese or Italian is spoken, usually also English. Although in such cases we do not make use of the vocabulary test in reckoning mental age, it had been given in 132 cases. The median vocabulary score for these children at each mental age is given in Table VI. For comparison the median scores for the three groups of American children for the same mental ages are repeated.

TABLE VI

	Below 86 IQ	86–115	Above 115 IQ	Latins
13–6 to 14–5		51.7	50	50
12–6 to 13–5	46.7	47.5	43.5	42.5
11–6 to 12–5	43.7	41.4	40.2	33.3
10–6 to 11–5	34	33.7	35.2	30
9–6 to 10–5	32.5	29	30.6	25
8–6 to 9–5	24	22.2	21.7	16.4
7–6 to 8–5	18.3	18	20	12.3
6–6 to 7–5	13.7	12.3	12.5	9

The fact that a majority of these children had learned another language before learning English is reflected in their inferior vocabulary scores for three or four years after entering school. After that, however, the vocabulary rapidly catches up with mental age. After the mental age of 12 years these children are practically on a par with their fellow pupils *of the same mental level* who have known no other language than English.

For the entire group of 132 Latins with mental ages ranging from 6 years to the "average adult" level, the correlation of vocabulary with mental age score was .84, or practically the same as for the American group with I. Q. below .86.

Sex Differences

Of the 631 children, 359 were boys and 272 were girls. As the group was a miscellaneous one no comparison of sexes could be made on the basis of chronological age, but the median vocabulary scores of boys and girls at each mental age were as follows:

Mental Age	Boys	Girls
18–6 to 19–5	75	76.2
17–6 to 18–5	73	70
16–6 to 17–5	65	68
15–6 to 16–5	61.6	61.8
14–6 to 15–5	51.7	56.5
13–6 to 14–5	50.5	51
12–6 to 13–5	45	49
11–6 to 12–5	42.5	40.5
10–6 to 11–5	35.7	32.5
9–6 to 10–5	31	31
8–6 to 9–5	24	21.8
7–6 to 8–5	20.7	17
6–6 to 7–5	13.6	12.5

The differences are so slight that they have little significance. The small advantages of first one then the other sex are probably due to the limited number at each mental age. On the whole, the boys are a shade superior up to 12 years, the girls thereafter. The increase with each sex, apart from irregularities, is fairly constant. Figure I shows that for the sexes taken together the curve of vocabulary growth by mental age is practically a straight line.

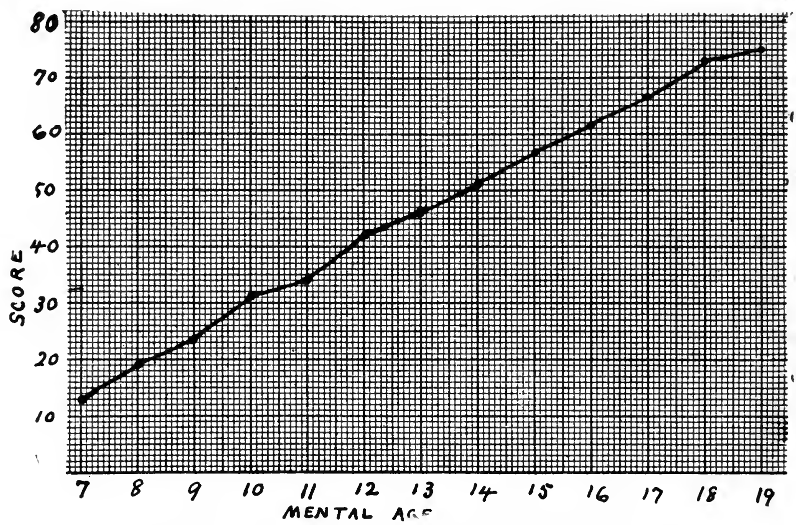


FIGURE 1. THE VOCABULARY SCORE AND MENTAL AGE

RELIABILITY OF VOCABULARY SAMPLING

One way to determine the reliability of a vocabulary test is to correlate the scores by the same individuals when tested by two or more different tests, each test constructed by the same method of random sampling from a dictionary. Five such lists were constructed by selecting every hundredth word from the Laird and Lee Vest Pocket Dictionary, the dictionary on which the Stanford vocabulary was based. The first list began with the first word in the dictionary, the second with the tenth word, the third with the twentieth, etc. As the dictionary contained approximately 18000 words, the selection gave 180 words for each list. All five tests were then given to 65 Stanford University students attending a class in education.

Most of the students were of the junior and senior classes. Mimeographed sheets were distributed containing the 180 words of one test, and the definitions were written. A period of 50 minutes was used for each test, one test being given each week. The laborious task of grading the written definitions was undertaken by Miss Mayme Anderson, a senior student in the department of education. The grading was done with extreme care. The names of the students were clipped from the papers so that the scores would not be influenced by any unconscious bias of the grader due to acquaintance with members of the class.

The intercorrelations of the tests were as follows:

Test	1	2	3	4	5	Corr. with av. of all
1		.83	.73	.78	.76	.92
2	.83		.80	.80	.78	.95
3	.73	.80		.80	.72	.90
4	.78	.80	.80		.76	.92
5	.76	.78	.72	.76		.84
Corr. with av. of all	.92	.95	.90	.92	.84	

Table VII, showing the correlation between test 3 and the average of all five, is typical.

TABLE VII

Score in Test 3

Average score in five tests	119–121	122–124	125–127	128–130	131–133	134–136	137–139	140–142	143–145	146–148	149–151	152–154	155–157	158–160	161–163	Total
158–160													1			1
155–157															1	1
152–154											3	1	1			5
149–151										2						2
146–148										3	1					4
143–145							1	1	3	1		1				7
140–142						1	4	4	3	3		1				16
137–139						1	4	1	1							7
134–136					1		3									4
131–133				1	1	1	1									4
128–130			1	4	1											6
125–127		1	1													2
122–124		1														1
119–121																
116–118																
113–115	2															2
Total	2	2	2	5	3	3	13	6	7	9	4	3	2		1	62

r = .90

The average of the correlations of five separate lists with the average score in all is .906. We can conclude, therefore, that a single list of 180 words is not greatly inferior in reliability to one of 900 words (that is, five of 180 words each.)

A better method of finding the reliability of a single test as compared with a composite of five is as follows:

(1) for a given test find each student's deviation from his average score in the five tests;

(2) find the probable error of these deviations.

These probable errors are as follows for the five tests:

Test 1,	1.6 words
Test 2,	2.0 "
Test 3,	1.5 "
Test 4,	1.7 "
Test 5,	1.8 "
Average P. E.	1.72 words

Since each word in the vocabulary list represents 100 words in the dictionary, the probable error of 1.72 in the score becomes 172 words in total vocabulary (as based on the dictionary used). In other words, the chances are even that a total vocabulary based on the 180 word list will not deviate more than 172 words from a total vocabulary based on a 900 word list. The chances are 6 to 1 that it will not deviate more than 354 words (2 P. E.); 25 to 1 that it will not deviate more than 516 words (3 P. E.); and 140 to 1 that it will not deviate more than 688 (4 P. E.).

For these 65 students the correlation was found between the average score on the five tests and the average class mark earned in all the university courses. As this was not done until two years after the test was given, the class marks were available for the four years of university work. The correlation was .28 Pearson.

Reliability of Vocabulary Lists with School Children

Two of the 180 word vocabulary lists used with college students were given by Mrs. Chamberlain to 32 school children who had been tested by the Stanford-Binet. The pupils were fairly evenly distributed from the first to the eighth grade. In this case the tests were given orally to the pupils taken one at a time as in a Binet test.

Each of these 180 word lists we have divided into three by selecting every third word. Thus we have six lists of 60 words each. These may be treated as six separate tests or they may be combined into tests of 120, 180, 240, 300 or 360 words. Table VIII gives the

intercorrelations of these six tests for the 32 children. The correlations were found by the Spearman Footrule and converted into values by the formula recommended for this purpose by Pearson.

TABLE VIII

	a	b	c	d	e	f	abc	def
a		.94	.93	.94	.94	.94	.98	.96
b	.94		.95	.94	.95	.92	.98	.98
c	.93	.95		.94	.94	.93	.97	.96
d	.94	.94	.94		.97	.95	.96	.99
e	.94	.95	.94	.97		.92	.96	.94
f	.94	.92	.93	.95	.92		.94	.97
abc	.98	.98	.97	.96	.90	.94		.98
def	.96	.98	.96	.99	.94	.97	.98	

The correlations in the above table are strikingly uniform and extremely high, decidedly higher than those found for the 180 word lists used with university students. This is not to be interpreted as indicating a higher reliability for a 60-word than for a 180-word list. The size of a correlation coefficient is greatly influenced by the heterogeneity of the subjects. In this respect the group of school children differed greatly from the university group. The latter represented a cross section of mental ability; the former ranged from six year mental age to "superior adult" by the Stanford-Binet scale.

It is a better measure of reliability of a 60-word list to find the P. E. of its deviation from the average of the six tests, as we have already done for the university students. Following are the P. E. values found for the deviations for each of the 60–word lists:

P. E. of Deviations

List a.......2.4 words
List b.......1.1 "
List c.......1.5 "
List d........1.5 "
List e.......1.2 "
List f........1.2 "

Average P. E. 1.48 words

Since each word in a 60-word test represents 300 words in the dictionary, the P. E. of 1.48 words in the score equals a P. E. of 444 words in total vocabulary. The chances are even that a total vocabulary (e.g., 7,200, 13,400, etc.) based on a 60-word test will not differ by more than 444 words from that which would result from a 360-word list. A P. E. of this amount is approximately 5 per cent. of the total vocabulary of an average 12 year child.

In like manner the average P. E. was computed for the three 120 word lists and found to be 2.13 words in score. Since each word of a 120-word list represents 150 words in the dictionary, the P. E. of 2.13 equals 320 words in the total vocabulary. The P. E. of a list of 180 words is 1.9 in terms of score, or 190 words in total vocabulary (since each word in this case represents 100 words in the dictionary).

The Selection of Crucial Words for Vocabulary Test

This will be made the subject of a separate report later. Correlation of each word of the Stanford vocabulary test with mental age shows that the words differ greatly in diagnostic value. Some present a very steep curve of per cent. passing at successive mental ages, others give a curve which is almost horizontal. The latter should of course give place to other words chosen for their diagnostic value. We believe it will be possible, before long, to measure the intelligence level almost as accurately by means of a vocabulary list of 100 crucial words as it can now be measured by any existent intelligence scale.

Vocabulary and the Test of Naming Words

In order to answer the question whether the ability to name a given number of words in three minutes is dependent upon the number of words known, we have correlated for a miscellaneous group of 360 children the scores earned on the Stanford Revision vocabulary test and the Binet test of naming words (three minutes). The correlation was .49, Pearson. The dependence is only moderately close.

Errors of Scoring

The writer has corrected the definitions, recorded verbatim, of 100 vocabulary tests made by three Stanford students who were being trained in Binet testing. The vocabulary scores in these tests ranged fairly evenly from 15 to 65, with an average of approximately 30. In only about 2 per cent. of the total scores given

the individual definitions had to be changed, and only in an insignificant number of cases did the errors in a test have any effect on the mental age earned. The criticism that the test is vitiated to any serious extent by difficulty of scoring certainly does not apply in the case of ordinarily competent examiners.*

However, there are certain words in the Stanford vocabulary list which bring out the personal equation in scoring more often than the others. Among them are *ramble, afloat, artless, outward, southern, noticeable, quake, nerve, sportive, peculiarity* and *selectman.* These should ultimately be replaced by others equally difficult but less troublesome to score.

Summary

1. For a miscellaneous group of 631 school children the correlation between vocabulary and mental age is .91.

2. The probable error of a mental age based upon the vocabulary test alone is only 9.6 months, in the case of school children, and the chances are approximately 6 to 1 that such a mental age will not be in error more than a year and a half. The probable error is only 12 months in the case of prisoners, hoboes, and other adults of widely different age, experience and schooling.

3. Children of a given mental age have approximately the same vocabulary regardless of chronological age.

4. Portuguese and Italian children from homes where a foreign language is spoken, are for the first two or three years of school life considerably below the median for American children of the same mental age. This difference, however, almost totally disappears by the time the child has attained the mental age of 12 years.

5. The median vocabulary at each mental age is practically the same for boys and girls.

6. Vocabulary growth is remarkably constant and regular, the curve of medians for the successive mental ages being almost a straight line.

---

*The following is a striking illustration both of the ease of scoring the vocabulary test and of its accuracy. A feeble-minded youth was brought to the writer for a mental examination. When the vocabulary test was being given the subject interrupted with the statement that he had read about "that word test" in the Literary Digest and that he had tested himself. Asked how many words he knew, he said 36. When the Binet test was finished it was found that the boy had defined 34 words correctly, which is exactly the median for 11 years. The mental age by the complete Stanford-Binet was also exactly 11 years. This feeble-minded boy had measured his own intelligence and missed it by only a third of a year!

7. When five different vocabulary tests of 180 words each were given to 65 university students the intercorrelations ranged from .72 to .83, with an average of .77. The average correlation of a single test with the total score for five tests was .906.

8. The probable error of the deviations of a single test from average score earned in five tests is only 1.72 words in the score, or 172 words in total vocabulary.

9. The correlation of the average score earned in the five tests with average university class work was .28.

10. Six vocabulary tests of 60 words each yielded in the case of 32 school children, intercorrelations of .92 to .98, with an average of .94.

11. The probable error of the deviations of a 60-word list from the average score earned in six such tests is only 1.48 words, or 444 words in terms of total vocabulary. The probable error of total vocabulary is only 320 words when the average of two such lists is taken, and only 190 for the average of three.

12. The correlation of vocabulary score (in 100 word list) with number of words named in three minutes (60 word test) was .49.

13. The errors in scoring the vocabulary test are negligible for examiners who have had adequate training.

COMMUNICATIONS AND DISCUSSIONS

A PRELIMINARY NOTE ON THE USE OF THE HAHN-LACKEY GEOGRAPHY SCALE

During the past few months the writer, with the assistance of some of his graduate students, has been able to make some studies of the results of the teaching of geography by means of the Hahn-Lackey scale. In one Brooklyn (N. Y.) school where 37 per cent. of the children are foreign born, various steps of the scale were given to 460 children in grades 4 to 8, inclusive. The average of all the tests given was 65.2 per cent. According to the scale the average should have been 74 per cent. The average of the results in the eighth grade in the school was within one per cent. of what the scale called for, using steps "S," "Q," "M," and "P" as a basis.

A second set of papers, based on the very simple steps "W" and "X" of the scale, has been collected in another large Brooklyn school with 3600 children. Of these children 35 per cent. are Hebrews and 10 per cent. Italians. The total number of children tested, grades 4 to 8 inclusive, was 1595. The grand average for all the children's papers was 74.06 per cent. According to the scale it should have been 94.2 per cent. Here, as in the first school cited, the children in the 8th grade made a better relative showing than those in the lower grades. They had an average of 90.1 per cent. where the scale called for 98.5. They were thus 12 per cent. more efficient in their answering of geography questions that they were supposed to know than were the 1595 children taken as a whole.

Comparing these two schools on the basis of steps in the scale which were entirely different, we find that the first school was 88 per cent. efficient, and the second 78.7 per cent. efficient in their results in geography.

A third set of papers based on steps "W" and "X" was collected in a summer school for children in a college town. 116 papers were rated in grades 4 to 7 inclusive. The average for all the children was 65.7 per cent. (Cf. 74.06 per cent., which was the average made by the second Brooklyn school on the same material.) The 116 children should have made 93 per cent. Expressed in percentage, this school is only 70 per cent. efficient in geography.

The "efficiency" rating for the three schools would be then, 88 for the first, 78.7 for the second and 70 per cent. for the small school.

Further investigations will be undertaken, and a full report published.

CHESTER A. MATHEWSON.

Brooklyn Training School for Teachers.

THREE METHODS OF TEACHING RADIO TELEGRAPHY

Last fall the Carnegie Institute of Technology started a night school course in radio telegraphy for drafted men of Class 1 in Pittsburgh. This offered an unusual opportunity for some experi-

mental work on the learning process of which we have availed ourselves. A two minute test in receiving the continental code with prescribed test material was given at every meeting of the class. Classes met three times a week. I shall report here one phase of this investigation, namely, the comparison between several methods of teaching the telegraphic code.

The classes in radio telegraphy which were organized November 19th and during the two following weeks were divided as nearly as possible into six groups of equal general ability on the basis of a series of mental tests. Two classes were taught by the visual method, two by the phonetic method, and two by the synthetic method.

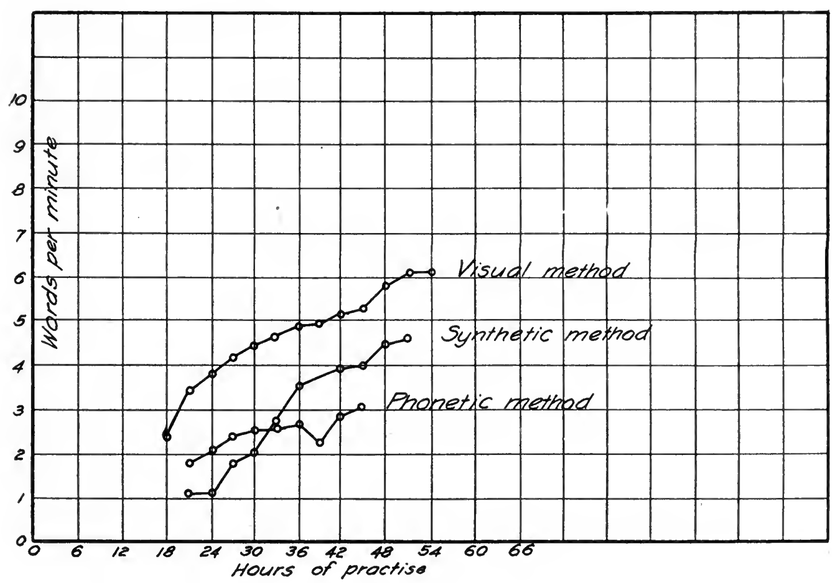
The Visual Method. The students were given cards on which were printed the dot-dash equivalents of the letters of the alphabet according to the continental telegraphic code. They were asked to write these dot-dash equivalents until thoroughly memorized. In class the students were given blackboard explanations of the dot-dash equivalents, but their practice in class was entirely devoted to receiving and interpreting the auditory impressions from a buzzer. The objection raised against this method is that the student thinks of dots and dashes instead of sounds and rhythm. The report reached us that a school of telegraphy in St. Louis emphasizes the sound of the letters, excluding all drill on the visual dots and dashes and that the learning time by this procedure is considerably lower than that of the visual method. We have tried this procedure under the name of the phonetic method.

The Phonetic Method. The students were asked to avoid learning the telegraphic code by any visual equivalents, and to concentrate their efforts on the sound patterns of the letters. We have good reason to believe that these instructions were quite generally complied with. Their entire practice-time was devoted to drill on the interpretation of the auditory sound patterns. They were probably not confused by thinking simultaneously about the visual and auditory equivalents of the alphabet. One serious objection against this method is that the student is deprived of the ready opportunity of studying the telegraphic code outside of class hours unless he can obtain practice by hearing others send. Practice by the visual method requires ony one man, while practice by the phonetic method requires, besides the student listener, a good sender. On this account the students who faithfully follow the phonetic method do not practice as many hours a week as the visual students. In spite of this obvious disadvantage the phonetic method was tried, since that is the only way to settle the question. The generally approved method of learning typewriting is the touch system by which the student is forbidden to look at the typewriter keyboard but this method is not the "obvious" one to the beginner nor to common sense.

During a discussion of these teaching methods in the seminar of the Department of Psychology at Carnegie Institute of Technology it was suggested that the analogy of the new methods of teaching reading in the public schools might be made use of. These methods

teach the student to differentiate between words and do not require him to analyze the individual letters. We decided to try this procedure and we have designated it the synthetic method.

The Synthetic Method. The student is taught to differentiate between several short words. The list of words is increased with practice and the student is expected to acquire the ability to recognize the individual letters not as individuals but as parts of words. This procedure produces remarkable results in teaching reading and it had not been tried in learning telegraphy as far as we could ascertain. One objection to the synthetic method is that the mortality in the night classes which were taught by the synthetic method was extraordinarily high. This was due to the discouragingly slow progress during the first few weeks which is to be expected with the synthetic method. Even if the method should be superior in the end this factor of initial discouragement must be reckoned with.



The Results of the Teaching Experiments. The accompanying diagram shows the progress of the three groups of radio students taught by the three methods. The curves are plotted with receiving speed in terms of words per minute as ordinates, and amount of practice in terms of *class* hours. The ordinates represent the average receiving speed of 38 students in the visual method, 51 students in the phonetic method, and 24 students in the synthetic method. These figures indicate that the mortality in the synthetic classes is greater than in the visual and phonetic classes.

The visual method is evidently superior to the phonetic and synthetic methods. I believe that the primary reason for this is that the visual method gives opportunity to practice the code outside of class hours without requiring the assistance of a second person. The curves shown here are not theoretically correct since the practice units of the curves are not comparable, but the curves do represent the practical state of affairs namely, (1) that the visual method affords opportunity for more practice hours per week than the phonetic and synthetic methods, and (2) that the advantage so gained by the visual method outweighs the advantages of the synthetic and phonetic methods. These conclusions do not necessarily hold for a full-time day course in telegraphy.

Recommendations. On the basis of observation of progress in these classes, we recommended the visual method for initiating the student to the telegraphic code. Practice outside of class hours may be by the visual method and is voluntary on the part of the student. Practice in class is limited mainly to drill in recognizing the sound patterns. As soon as the code has been mastered, word drill is introduced to help the student in recognizing words as such rather than by individual letters. The addition of a series of two hundred sample telegrams furnished us through the courtesy of Col. J. B. Allison, of the Signal Corps, has been a vital factor in keeping the interest of night students in the rather monotonous code practice.

L. L. THURSTONE.
Division of Applied Psychology,
Carnegie Institute of Technology, Pittsburgh, Pa.

The Journal of Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD.

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

NOTICE TO SUBSCRIBERS. Notify the publishers, Warwick & York, Inc., of any change of address; as copies lost in transit because of subscriber's removal to another address will not be duplicated.

THE JOURNAL is issuing a series of Monograph Supplements for the publicaiion of more extensive researches in educational psychology. For details prospective contributors may apply to H. E. Buchholz, 10 East Center St., Baltimore, Md.

The price of the JOURNAL is $3.00 per year of ten monthly issues, payable in advance. Single issues are 50 cents each. Back volumes are $5.00 in original wrappers and $6.25 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME IX OCTOBER, 1918 NUMBER 8

EDITORIAL

EXPERIMENTAL STUDIES IN EDUCATION

In an address at Atlantic City last spring Professor Judd said that one reason why we do not hear so much about educational measurements now as we did a year or two ago is that they are becoming a part of the routine work of the school. Some may deny the underlying assumption, and insist that we now hear much more about them than ever before,—that there never has been a time when there was such activity in the construction, perfection and application of educational scales as at present. The contest of educational measurements for recognition has pasek. No progressive, alert, superintendent, principal or teacher denies their theoretical significance and possible value. But whether they have generally become a part of the routine work of the school may well be questioned. Would that it were so. While most teachers admit that educational measurements are good things (when applied in some other school), and while many have a vague feeling that the experimental movement is full of promise, it is precisely the systematic use of measurements in the routine

work of the school that is conspicuously lacking. All too generally they are looked upon by the teacher as a new sort of club held over her by some outside power (the superintendent or a survey commission) to drive her to more strenuous exertions. Rarely has the teacher come to see that they can be used as aids to enable her to accomplish her purposes with more certainty and less expenditure of effort.

To make educational measurements better known to teachers and school officers, to foster the experimental attitude in education, and to devise ways and means of employing educational scales in school routine, are the aims of a new local organization in New York City known as the New York Society for the Experimental Study of Education. The society meets one evening a month during the school year for reports on experimental investigations in progress, discussions and plans for new investigations, and abstracts of experimental studies carried on elsewhere. The membership includes superintendents, high and elementary school principals and teachers, officials of bureaus of educational research and educational foundations, and representatives from teacher training schools, colleges and universities.

It is to be hoped that the society will prove a rallying point for all those interested in the scientific study of education in that locality, and will contribute materially to a wider appreciation of the value of such study to the school administrator, to the teacher and to the pupil. Particularly in view of the social and economic reconstruction which impends after the war there is need for educational authorities to work out methods to determine quickly and with a fair degree of accuracy the individual peculiarities and abilities of each child, to decide upon the kind of educational activities that will best fit him for his probable life work, to ensure adequate opportunity for the development of those broader interests which are necessary to good citizenship and a cultured personality, and to individualize instruction so that he will always be confronted with tasks adjusted to his powers and attainments. These ends can only be realized as the result of painstaking and carefully controlled studies carried on co-operatively by school people under school conditions. The society has the hearty approval of Dr. W. L. Ettinger, the recently elected superintendent of schools, and gives promise of great influence and usefulness. J. C. B.

NOTES AND NEWS

The latest price list and circular of information from the Bureau of Educational Measurements and Standards of the Kansas State Normal School at Emporia contains answers to twenty questions frequently asked concerning educational measurements, a brief description of each of the more important standardized tests, and a bibliography of the more significant articles in the domain of each school subject.

The Indiana University Bureau of Cooperative Research, under the new director, Dr. Walter S. Monroe, announces a series of standardized reasoning tests in arithmetic. The present arrangement of the tests is based on returns from over 13,000 pupils collected last spring. The tests are arranged in three sets, one for grades IV and V, one for grades VI and VII, and one for grade VIII. A final standardization of the tests is planned for the current year.

A committee on standard tests and measurements of the Connecticut School Superintendents' Association, consisting of Ernest C. Witham, Southington, and Carlon E. Wheeler, New London, has recently published an interesting report. A schedule is presented providing for one or more standard tests in each month in the school year. In this way the burden of testing is distributed, and the administration of tests is made a routine matter. The report contains a brief discussion of the most valuable tests in each school subject.

The Municipal Reference Library, City of New York, Dorsey W. Hyde, Jr., librarian, has recently inaugurated a series of special reports on civic subjects. The first number contains a list of references on "What to Read on New York City Government," and the second presents an interesting discussion of "Teaching Citizenship via the Movies."

The Child Health Organization of New York City is making an effort to direct the attention of teachers to the physical welfare of their pupils. It has adopted the slogan of "Scales in every school," and an effort is being made to get teachers to keep a chart of each pupil's height and weight each week, and thus keep a careful watch over the pupils' nutrition.

The Public Charities Association of Pennsylvania is endeavoring to facilitate the formation of special classes in the public schools of every community in the state for the special instruction of men-

tally exceptional children. It has prepared a bill for the next legislature providing state aid for school districts which organize such classes. It is gathering authentic data as to the precise extent of mental retardation and deficiency among children of school age, and through mental clinics and examining stations it is trying to aid in the identification and classification of backward children. This work is largely under the supervision of Dr. Norbert J. Melville, of the Philadelphia School of Pedagogy.

Dr. Harry Kirke Wolfe, professor of philosophy in the University of Nebraska, died on July 30 at the age of fifty-nine years. Dr. Wolfe was one of the pioneers in experimental psychology and its applications to education in America.

Dr. Eleanor H. Rowland, formerly dean of women at Reed College, has recently become head aide at the Walter Reed Hospital, in Washington, D. C., where the first returned American soldiers have been received.

Professor C. E. Seashore, of the University of Iowa, is conducting investigations on certain problems of hearing as related to the army and navy, and is also devising and standardizing a series of tests for the selection of telegraphers and radio operators. R. H. Sylvester, assistant professor of psychology in the same university, is now lieutenant and chief clinical psychologist at Camp Dodge.

Dr. H. L. Hollingworth, associate professor of psychology in Barnard College, has been commissioned captain in the sanitary corps.

Dr. Samuel C. Kohs has been elected assistant professor of psychology at Reed College, Portland, Oregon.

Dr. Cyrus D. Mead, who for six years has been assistant professor of elementary education at the University of Cincinnati, has been appointed associate professor in the University of California. Before leaving Cincinnati Dr. Mead was elected president of the Cincinnati Schoolmasters Club.

At Harvard University Dr. Herbert Sidney Langfeld has been appointed acting director of the psychologcal laboratory.

At the University of Michigan Associate Professor J. F. Shepard has been made professor of psychology, and Dr. H. Foster Adams has been advanced from instructor to assistant professor of psychology.

VOLUME IX NOVEMBER, 1918 NUMBER 9

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

THE DIAGNOSTIC VALUE OF THE WOODY ARITHMETIC SCALES: A REPLY. PART I

W. W. THEISEN AND CECILE WHITE FLEMMING

Bureau of Educational Measurements, State Department of Public Instruction, Madison, Wisconsin

In the issue of *School and Society* for October 6, 1917, an article appears by Dr. Walter S. Monroe entitled "An Experimental and Analytical Study of Woody's Arithmetic Scales." In view of the extended use attained by the Woody scales during their first year and a half, this title is unusually attractive. The author of the article states that "It is the purpose of this study to learn how completely it is possible for these scales to diagnose a class, and how accurate is such diagnosis as they make." The article is characterized by an absence of any thoroughgoing attempt to show what merits the tests possess as diagnosing instruments and by an effort to prove the author's main contention that they do not diagnose through attempting to show in a rather superficial way that the Cleveland Survey tests do diagnose. The article does a grave injustice to both series of tests and to the cause of educational measurements, because it fails to reveal the real diagnostic powers of either of the tests. An untrained reader might very readily conclude that neither test possessed much diagnostic merit.

It is not the purpose of the present writers to attack the statement that the Cleveland Survey tests can be used to diagnose. In fact the great number of examples of various types which they contain make them quite comprehensive and useful for the purpose whenever sufficient time is available for examining fully the symptoms revealed. We shall however in the first part of this paper attempt to answer the principal criticisms which Dr. Monroe has made of the Woody tests. In the second we shall endeavor to point out some of the possibilities of the tests as diagnosing instruments.

I

In the first place, to diagnose means more than to find class scores and to compare them with established standards. The article referred to above treats diagnosis in this limited sense. So long as a diagnosis is only in terms of standards of attainment, it is superficial. It must go further. It must reveal the particular processes in which pupils fail, not necessarily the examples. It must reveal the particular mental processes or habits at work which are the seat of the difficulty. When a third grade child adds example 3 of Woody's addition test $\begin{array}{r}17\\2\\\hline\end{array}$ and obtains 10 as the answer, it is of little service to treat this merely as an error. The fact to be considered is that the child has added two, seven, and one, and obtained ten. This habit must be broken down and the habit of two-column addition built up before the child can be expected to progress satisfactorily. In any case a diagnosis must seek to reveal the mental processes at work tending to produce failure. Hence we must not be led astray by assuming that Dr. Monroe's exhibit of scores made on the Woody and the Cleveland Survey Tests represent the diagnosing merits of either test. Some conception of the diagnosing merits of the Cleveland Survey Tests may be had by reference to Counts' study.*

Monroe contends quite correctly that "arithmetical ability" consists not of a single ability, but several abilities. He points out for instance that Ballou records 14 types of examples in the addition of two fractions. This makes way for his first criticism of Woody's tests, namely, that they do not include all important types of examples required to test these several abilities. In reply it may be said that while the greater number of types may be admitted, it does not necessarily follow that children must be tested on all of them before fundamental weaknesses in the work of a class can be discovered. The tests do not contain all types and it is not essential to a working diagnosis that this be the case. A working diagnosis must furnish clues as to the most probable causes of failure—clues whch can be used as starting points for remedial measures. This the Woody tests do. When a teacher discovers that a number of pupils in attempting to add fractions where the

*George S. Counts. *Arithmetic Tests and Studies in the Psychology of Arithmetic,* University of Chicago Press, Supplementary Educational Monographs No. 4, 1917.

addition is indicated by the + sign, add numerators to numerators, and denominators to denominators, or that some add numerators to denominators, it is not necessary to extend the testing to all types of examples in fractions. In one small class of nine sixth grade pupils tested by Woody's Addition Scale, Series A, six added numerators to numerators and denominators to denominators, in each of the six examples in addition of fractions where the addition is indicated by the plus sign, *e.g.*, in example 27, $1/8 + 1/4 + 1/2 =$, the answer $3/14$ was obtained through such a process. In this school the course of study calls for addition of fractions, including the types indicated by these examples, in grade five. The results of the test are good evidence that the subject was not well taught, and the prevailing type of error indicates the nature of the remedy.

Furthermore, we must not lose sight of the fact that a satisfactory test must have commendable administrative features as well as supervisory and psychological values. While a test must afford a working diagnosis of teaching needs, at the same time it must be such that it can be readily administered. For that reason it cannot contain an unduly large number of examples. The inference which the writer leaves us to draw, namely, that the Cleveland Survey tests provide more types of examples than do the Woody, is not founded upon fact, as one may see by a careful comparison of the two series of tests. All things considered, Woody has combined the essentials of a good test and scale to a remarkable degree. To meet all the requirements that Mr. Monroe is asking would necessitate tests of unwarranted length and tests which would require more labor to summarize properly and intepret than most school systems can afford. The four Woody tests require eighty minutes as it is.

The criticism of the tests on the ground that the diagnosing power of a test is limited to the types it contains is rather unjust. By following up the test with slightly different examples of about the same degree of difficulty, one may discover other weaknesses than that expected from a given type of example itself. Furthermore, in all testing a basic assumption is that the examples or questions selected represent a random selection of the entire field. If correlation studies have shown us anything, they afford grounds for believing that if pupils are weak in particular types of examples, they are very apt to be weak in types somewhat similar. Hence

when wrong or confused mental habits are discovered, it is not necessary to test on all types before we can predict certain types on which they would probably fail.

Dr. Monroe makes the statement that "a satisfactory diagnosis must either be complete, or, if not, its limits must be known. Otherwise the mistake will be made of assuming that a partial diagnosis is complete." In reply we may say first, that we know something of the limitations of the test when we know that it does not contain all types or ramifications of types of examples. Nowhere does Woody claim that his tests will furnish a complete diagnosis. He does say that they will "aid greatly in locating the weaknesses of the class,"* a statement not out of keeping with the truth. Second, a satisfactory diagnosis need not necessarily be complete in all details. A diagnosis is satisfactory which reveals a number of specific class weaknesses or incorrect habits formed which are causing errors. The teacher desires to know where to attack in order to bring about improvements. She does not attempt to remedy all weaknesses at once. As we have said the most serious of them furnish her with starting points. If she be at all competent she will make use of additional examples somewhat similar in nature, and difficulty in order to discover other limitations of the class.

For the purposes of diagnosis it is not so important to know how many errors were made on particular examples as it is to know what is causing the errors. There is a vast difference between errors as such and recurring errors of a similar nature, caused by the same mental processes or habit. This is a distinction which the writer of the article apparently disregards when he discusses the diagnostic features of tests. When several pupils in a class give the same incorrect answer to an example, or when several errors of a similar nature are made by an individual, the difficulty can usually be traced to a single source, *e.g.*, in answer to example 16 of the subtraction test $\begin{array}{r}270\\190\\\hline\end{array}$ it was found that 120 was given by a number of children in several classes in one city. How to account for it was for a time a puzzle. It finally developed that it was probably due to an insufficient familiarity with the Austrian method of subtraction. The mental processes involved after putting down the cipher was somewhat as follows: "Seven and what number

*Clifford Woody. *Measurements of Some Achievements in Arithmetic*, p. 23.

make nine? Put down two. One and what number make two? Put down one." Example 18 of the subtraction scale $\begin{array}{r}1000\\537\\\hline\end{array}$ was usually solved similarly by the same pupils. It was discovered that these children in the second and third grades had been taught to subtract by the common method of borrowing The present teachers of these children, now in grades four and five, use the Austrian method, teaching subtraction as a form of additon. Hence the confusion. Now to have treated the answer 120 to example 16 merely as an error would not be diagnosing an evident need of this class. To discover the mental process back of the error is to diagnose. The number of such mental processes causing errors which can be discovered through the use of the Woody tests is sufficient to furnish reasonably satisfactory diagnosis of teaching needs.

A third objection is raised on the ground, that Woody, as Monroe claims, has used the statistical rather than the analytical method of selecting his examples. As a matter of fact, Woody has used a combination of the two methods. We wrote to Dr. Woody for a definite statement on this point. In a reply he says "The problems for my scales were chosen primarily for the purpose of analysis of errors made by pupils."* Statistical considerations were secondary. Nearly every example was selected because it added some new feature not contained in others already chosen. The examples rejected because of their failure to meet certain statistical requirements were few as any one may discover by correspondence with Dr. Woody himself. In this letter he states that "Those problems which were rejected had a particular characteristic which made them more or less local in their nature. By local I mean such a peculiarity that a class in the third grade would understand the meaning of such a problem while a class in the fourth would not." He says further with reference to the statistical selection of examples in Series "A"—"The steps were not equal between the different problems. This is due to the fact that the problems were arranged in order of difficulty and given that particular order because I did not wish to throw out problems merely from a statistical point of view. I preferred to retain the problems because of their diagnostic value rather than throw them out because they did not make uniform steps which a statistical scale would be inclined to do."

*Letter of January 23, 1918.

One decided advantage gained when we select examples in order of increasing difficulty, as we do when we use the statistical method, is that we obtain a graded series. The advantage of such a series in determining achievement in arithmetical abilities is one which should not be overlooked. This feature of Woody's scales resulting from the use of the statistical method is one for which no other set of tests in fundamental operations provides to a like degree, and it is one whose value Dr. Monroe fails to consider at all in his criticism of the statistical method of selecting examples.

Another objection to the statistical method of selecting examples is raised by the writer of the article because only those examples were chosen which were solved by a gradually increasing percentage of pupils from the lower to the higher grades. The writer states that "Under normal conditions the per cent. of pupils doing the examples correctly should not increase gradually from the lower to the higher grades." He does not take into account the fact that a gradual increase in the sense used by Woody means that *beginning with the grades in which some of the pupils are able to solve a given example there should be a gradual increase* in the percentage of pupils in each grade above who solve it correctly and *not a decrease* from one grade to the next.

Moreover, Woody attached no weight to scores made on examples solved by less than 2.15 per cent. of the children in the grade (*i.e.*, at a distance greater than 3 P. E. from the median of the distribution for the grade) in locating the examples on the scales. Examples solved by more than 2.15 per cent. but less than 25 per cent. (*i.e.*, between 1 P. E. and 3 P. E. from the median) were given only half weight. Thus no one need be concerned about the importance attached to examples much too difficult for children of the grade in making up the scales.

The writer of the article states that "Before a type of example is taught a pupil can not be expected to possess the ability to do it." Nevertheless, much to our surprise, the child sometimes does. But however that may be, we know of no one who contends that he should. Certainly Woody does not, for nowhere in his monograph do we find a statement to the effect that he does. Monroe says further that "at the time in his school career when it is taught he will learn more rapidly than afterwards." He refers to the median achievement of pupils on the Cleveland Survey Tests as tending

to support his statements. For the benefit of the reader the table showing these scores is reproduced here, Table I. Dr. Monroe calls particular attention to the scores on examples in fractions.

Tests H and O as used in the Cleveland Survey are tests in fractions. These figures indicate no score for grades 3 and 4 on test H and for grades 3, 4, and 5 on test O. Evidently the type of example included in test H was first taught in grade 5 and that of test O in grade 6. Obviously the changes in score from 0 in grade four to 5 in grade five on Test H, and from 0 in grade five to 3.1 in grade six on Test O, appear large. But this change can not well be interpreted as a "*gain from grades four to five*" *much larger than the improvement between subsequent grades.* The figures indicate rather a change from a condition of *no ability* before the examples have been taught to a condition of *ability* resulting from the teaching. After this ability appears there is an uninterrupted increase throughout the succeeding grades for both Test H and Test O. Moreover Table I shows, for nearly all of the other tests, a similar regularity of improvement from grade to grade. This may be seen graphically for a number of the tests from Figures I and II. The curves for tests E, F and I, shown in Figure I, represent tests in addition,

TABLE I

*Medians in Each Arithmetic Test for All Grades—Cleveland Survey**

	Grade					
Test	3	4	5	6	7	8
A	13.4	17.8	22.2	24.8	26.7	27.5
B	9.3	13.4	17.2	19.8	21.5	26.0
C	6.5	12.0	15.5	16.6	17.7	19.0
D	6.3	12.4	15.7	18.5	20.8	22.5
E	4.3	5.3	6.3	6.8	7.5	7.8
F	2.0	4.9	6.7	7.5	8.6	10.1
G	2.0	3.9	5.2	5.5	5.9	6.6
H	0.0	0.0	5.0	5.5	7.7	8.5
I	0.6	1.1	2.0	3.1	4.0	4.7
J	1.9	3.2	4.0	4.4	4.9	5.7
K	0.0	4.0	6.8	8.5	10.1	12.5
L	0.0	1.7	2.5	2.8	3.2	3.9
M	1.4	2.5	3.2	3.8	4.4	5.1
N	0.0	0.8	1.3	1.7	2.0	2.6
O	0.0	0.0	0.0	3.1	4.1	5.5

*From CHARLES H. JUDD: *Measuring the Work of the Schools*, Table X, page 98.

subtraction and short division respectively. Test E is composed of single column addition examples of five numbers. Test F includes examples of the type $\begin{array}{r}1269\\ 772\\ \hline\end{array}$, and Test I of the type $6\overline{)57432}$. In none of these is there anything but a gradual improvement from grade to grade. Moreover, were one at first thought inclined to accept Monroe's assertion that "at the time in his school career when it is taught (*i.e.*, a given type of example) he will learn more rapidly than afterwards," he would have evidence from Test F that the statement is far from true. The median score for the fourth grade is more than double that of grade three.

In Figure II, tests L, N and O are represented. Test L is a test in multiplication of four place numbers by two place numbers. Test N is a test in division of the type $48\overline{)28464}$, and test O is a test in fractions. The curves for these tests confirm the evidence of gradual improvement shown in figure I. Certainly the facts of Table I from the Cleveland Survey Tests cited by Dr. Monroe do not offer satisfactory evidence to support his objection to Dr. Woody's statistical method of selecting examples. Still another objection is offered to the Woody tests on the ground that too few of the fundamental combinations are included. He asserts that the addition test, for example, includes only two examples in fundamental combinations. The addition scale has in effect not only two examples, but six that offer pupils an opportunity to exhibit their knowledge of the fundamental combinations. All but number two of the first seven examples, call for fundamental combinations. As may be seen they are additions of two numbers without carrying.

$$\begin{array}{cccccc} (1) & (3) & (4) & (5) & (6) & (7) \\ \begin{array}{r}2\\3\\ \hline\end{array} & \begin{array}{r}17\\2\\ \hline\end{array} & \begin{array}{r}53\\45\\ \hline\end{array} & \begin{array}{r}72\\26\\ \hline\end{array} & \begin{array}{r}60\\37\\ \hline\end{array} & 3+1= \end{array}$$

Though an example may be of two columns, it serves to diagnose ability in the fundamental combinations so long as carrying is not required, *e.g.*, example 5, $\begin{array}{r}72\\ 26\\ \hline\end{array}$ gives a child an opportunity to exhibit his knowledge of the combination 6 and 2, and that of 2 and 7. In the subtraction test there are ten examples, numbers 1, 2, 3, 4, 5, 8, 9, 10, 11, 12, that are essentially fundamental combinations. These are:

$$\begin{array}{cccccccccc} \begin{array}{r}8\\5\\ \hline\end{array} & \begin{array}{r}6\\0\\ \hline\end{array} & \begin{array}{r}2\\1\\ \hline\end{array} & \begin{array}{r}9\\3\\ \hline\end{array} & \begin{array}{r}4\\4\\ \hline\end{array} & \begin{array}{r}59\\12\\ \hline\end{array} & \begin{array}{r}78\\37\\ \hline\end{array} & 7-4= & \begin{array}{r}76\\60\\ \hline\end{array} & \begin{array}{r}27\\3\\ \hline\end{array} \end{array}$$

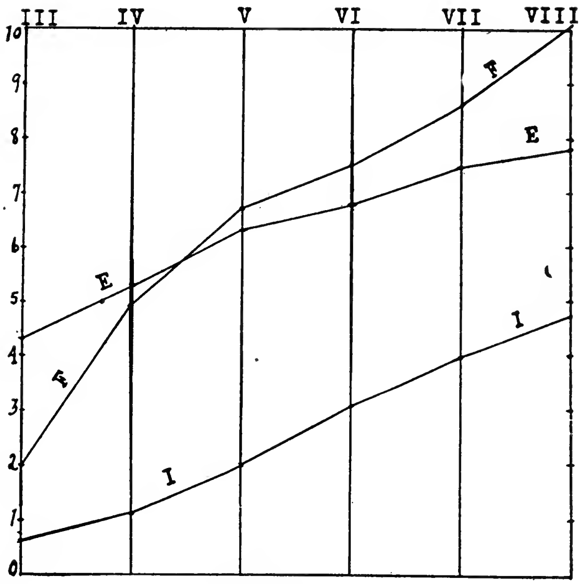


FIGURE I

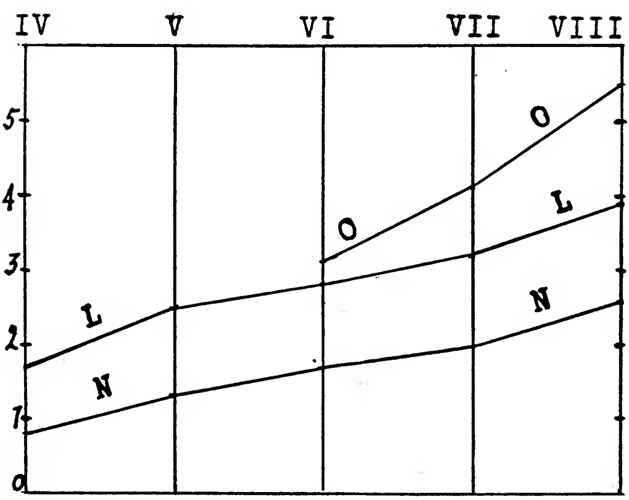


FIGURE II

In the multiplication test the first eight examples are serviceable for measuring ability in multiplication without carrying.

$3\times7=\qquad 5\times1=\qquad 2\times3=\qquad 4\times8=\qquad \begin{array}{r}23\\3\\\hline\end{array}\quad \begin{array}{r}310\\4\\\hline\end{array}\quad 7\times9=\quad \begin{array}{r}50\\3\\\hline\end{array}$

Eleven of the first twelve division examples are of the simplest kind. They include division by 1, 2, 3, 4, 6 and 9 without remainders or carrying.

$3\overline{)6}\quad 9\overline{)27}\quad 4\overline{)28}\quad 1\overline{)5}\quad 9\overline{)36}\quad 3\overline{)39}\qquad 4\div2=\quad 9\overline{)0}\quad 1\overline{)1}$

$6\times\ldots\ldots=30\qquad 2\div2=$

It will thus be seen that the tests offer many more opportunities for diagnosing abilities in fundamental combinations than appear at first glance. The truth is that the tests and the performances upon them require careful study to discover all that they can reveal. Not many examples could be added to the tests in their present form without necessitating the elimination of others. Monroe in demanding numerous additions to the tests strangely enough fails to indicate any desirable eliminations. To increase materially the number of examples on the test would as we have pointed out elsewhere make the use of the tests a very laborious and burdensome process.

The writer of the article bases his objection to the test of ability in fundamental combinations made by the Woody addition scale on the ground that pupils who made a perfect score on the example $\begin{array}{r}2\\3\\\hline\end{array}$ made less than average scores on the Cleveland Survey Test A, consisting of examples in fundamental combinations. But what did these same children do on examples 3, 4, 5, 6, and 7 of the Woody addition test, we ask? Were they able to make all of the fundamental combinations in these as well? Monroe states "that the Woody Scale revealed no differences in the ability of these pupils in using the addition combinations. The Cleveland Survey Test A reveals quite striking differences in abilities." But is it true that there were no differences? Where are the facts on examples 3, 4, 5, 6 and 7?

Dr. Monroe raises the further objection that the addition test

contains only one example of the type of example 18 which is He reports a ninety-three per cent. score on this example by a seventh grade class which when tested by the Cleveland Survey Test M scores below average. Woody has at least six other examples of two or more columns in more than two numbers, requiring carrying, as follows:

2563
1387
4954
2065

(11)	(13)	(15)	(16)	(17)	(22)
32	23	100	9	199	547
59	25	33	24	194	197
17	16	45	12	295	685
		201	15	156	678
		46	19		456
					393
					525
					240
					152

The critic fails to record what this class did on these examples. The same type of ability required in the first three columns of example 18 is represented in example 17. These two examples together with example 15 would give a more accurate idea of the ability of the class to add and carry than 18 alone. To the examples above might be added examples 19, 20, 21, 24, 31 and 33, which, though they involve decimals and United States money, may also be used to diagnose ability in addition. These examples are:

(19)	(20)	(21)	(24)	(31)	(33)
$ .75	$12.50	$8.00	4.0125	113.46	.49
1.25	16.75	5.75	1.5907	49.6097	.28
.49	15.75	2.33	4.10	19.9	.63
		4.16	8.673	9.87	.95
		.94		.0086	1.69
		6.32		18.253	.22
				6.04	.33
					.36
					1.01
					.56
					.88
					.75
					.56
					1.10
					.18
					.56

Apparently the writer of the article has singled out a few examples and recorded not causes of error but rights and wrongs on these few examples, forgetting that there are others in the test. He would have us use the Cleveland Survey Tests which contain but a very few types of examples in addition. They contain one test in simple combinations, two in single column addition, one in addition of similar fractions, one in addition of examples similar to Woody's 18th, and one containing examples in addition of dissimilar fractions. They include no examples in decimals, U. S. money, denominate numbers, or mixed numbers. Without any thoroughgoing attempt to reveal the diagnosing powers of either test he assumes to relegate one test to the scrapheap and to elevate the other upon a pedestal.

Finally, he points to the subtraction combinations on the Woody Scale, nine in number, and to examples 16, 17 and 18, which are:

270	393	1000
190	178	537

as being insufficient to reveal weaknesses in these phases of the arithmetical work. In fact he tries to show that the results attained with them may be materially different from those obtained on similar but longer tests. To substantiate this position he resorts to a comparison of performances made by a single class from each of grades 4 to 8 inclusive on these examples and on the Cleveland Survey Tests B and F which consist of similar examples. In the first place the number of pupils tested was too few to warrant any safe conclusions. Second, he records the average *per cent.* of examples correct on those from the Woody Scale and the average *number* of examples correct on the Cleveland tests. The latter he compares with the standard average on the Cleveland tests, but he fails to record the standard per cent. of examples correct for these examples of the Woody Scale. Such computations obtained by entirely different methods are not comparable. In order to make any sort of comparisons between the two tests from this standpoint, the standard per cent. of examples correct should be calculated. Then it will be possible to say that a given class was below standard as judged by one test or above as judged by the other. From Dr. Monroe's statements one is led to believe that children who scored high on the Woody examples did not do well on the Cleveland tests. For he says of the subtraction combinations, "According to

the Woody Scale the pupils do these examples very satisfactorily * * * * according to the Cleveland Survey Test these pupils are below standard excepting in the sixth grade, and there they are just up to standard." Again he says of examples 16, 17 and 18: "It should be noted that according to Test F the seventh and eighth grade classes are below standard in ability to do this type of example. The Woody scale shows these pupils to stand very high in ability to do examples of this type."

Now are the impressions intended here in keeping with the facts? Table II gives the average and standard number of examples correct on the Cleveland Test B and the average and standard per cent. correct on the examples from the Woody Scale. The last of these are computed from the tables in Woody's monograph.* The others are taken from the tables given by Monroe.

TABLE II

Showing the Average and Standard Number of Examples Correct on the Cleveland Survey Test B and the Average and Standard Percent of Examples Correct for Examples 1, 2, 3, 4, 5, 6, 7, 10 and 13 of the Woody Subtraction Scale

	IV	V	VI	VII	VIII
Average Cleveland Survey Test B....	10.6	12.7	19.3	18.3	18.
Standard average survey test B......	11.3	16.	18.3	21.4	24.4
Average per cent of examples correct Woody Scale....................	96.	99.	99.	97.	96.
Standard per cent of examples correct Woody Scale....................	92.9	97.4	98.6	98.8	99.4

TABLE III

Showing Facts Similar to Table II for the Cleveland Survey Test F and for Examples 16, 17, 18 of the Woody Subtraction Scale

	IV	V	VI	VII	VIII
Average Cleveland Survey Test F....	1.6	4.9	7.4	7.7	7.2
Standard average survey test F......	3.9	6.4	7.3	9.	10.2
Average per cent correct Woody Scale	54.	76.	92.	86.	90
Standard per cent correct Woody Scale	49.	79.	87.	93.	95.

*Ibid., Tables XI and XXII. The order and numbering of examples in Table XI is that used by Woody in collecting the original data and differs from the present order on the scale. Table XXII furnishes a key to the corresponding orders.

It is to be observed from Table II that when measured by either test grade six is above standard and grades seven and eight are below. Thus for these grades the two tests do not point to different conclusions. According to Table III,* grade six is again above standard on both tests and grades five, seven and eight are below standard on both tests. This is far different from what the writer would have us believe. Even from his own figures, derived as they are from but one class in each grade, the tests are shown to point in the same direction. The assertion made that "The Woody Scale shows these pupils (*i.e.*, grades seven and eight) to stand very high in ability to do examples of this type" does not accord with the facts. The two tests do diverge in grades four and five of Table II and in grade four of Table III. Two possible explanations may be offered. Tests given to more than 7000 Wisconsin children indicate that Woody's standards for each of his tests are too low in grades four and five. Standards in both tests probably need refining. Then the element of speed may be more effective in these grades. The averages on the Cleveland tests are based upon scores made in thirty seconds. They do not represent the ability to do examples of a given type, but the ability to do a given number of them in thirty seconds, two quite different things. Furthermore, one might well question the validity of results obtained from a limited number of pupils in so short a time. An error of ten per cent. in either direction in the time actually taken by a class is not impossible or unusual. From the facts which Monroe has given us we must conclude that the writer of the article has failed to substantiate his claim that the Woody test in subtraction contains an insufficient number of examples of the type just discussed, and that conclusions drawn from it may be misleading.

*The standard per cent. correct for the Woody examples is computed from Table XXIV of Woody's monograph. See his Table III column headed "Subtraction," and Table XXV column "Final Value," for the key to corresponding orders.

(*Part II will appear in the December number.*)

CORRELATIONS OF IMMEDIATE AND DELAYED RECALL

ARTHUR I. GATES

Teachers College, Columbia University

A considerable number of studies have accumulated bearing upon the general relation of "speed of learning" and "retentiveness," yet authorities disagree widely in their interpretation of these results. The terms "speed of learning," "retentiveness" and the like give scarcely an accurate description of the facts, since "speed of learning" has been determined in nearly all cases by an objective measure of the amount of material recalled at a certain time, viz., immediately after a study period and likewise "retentiveness" is determined by a similar measure of amount recalled at a particular later time. It will be seen that the terms "immediate recall" and "delayed recall" give a more accurate description of the functions concerned in the present experiment.

With regard to interpretation of previous studies, we find in current textbooks such conflicting statements as these:

"The one who learns quickly, ... retains it just as long and on the average longer than the one who learns more slowly . . . . Quickly come, slowly go."*

Pillsbury takes a middle ground:

"Individuals who learn easily seem to forget slowly while those who learn slowly forget rapidly. This law holds if one considers pure rote learning. When learning sense materia by log cal connections, the man who learns slowly may have an advantage if he gives the added time to understanding the matter. In this case the evidence shows that slow learning is compensated for by retentive memory."†

At the other extreme is Meumann, from whom pertinent statements are quoted. "Learners may, in general, be classified into two groups; rapid learners and slow learners . . . . . each represents a characteristic mental type. The typical differences persist no matter whether the material to be learned is of significant or a meaningless sort . . . . The typically rapid learner is usually a rapid forgetter. The rapid learner possesses no guarantee of permanent retention. The profit which the slow

*Strayer and Norsworthy. *How to Teach,* N. Y., 1917. p. 81.

†W. B. Pillsbury. *Essentials of Psychology,* N. Y., 1914. p. 198.

learner derives from practice is, in most cases, relatively much greater. The rapid learner is always found to be at a disadvantage when called upon to reproduce freely without re-learning."*

These statements of Meumann are an illustration of the doctrine of compensation which have been widely voiced by German psychologists. The belief in negative correlations between desirable traits, as the quotation well shows, has usually gone hand in hand with the attempt to classify individuals into types.

Thorndike, in his survey of correlational psychology, has vigorously assailed this doctrine:

"It should be noted that in original nature the rule is correlation, not compensation . . . . . It is very, very hard to find any case of a negative correlation between desirable mental functions. Divergences toward what we vaguely call better adaptation to the world in any respect seem to be positively related to better adaptation in all or nearly all respects. And this seems especially true of the relations between original capacities."†

It is in connection with the relation of learning and retention, as measured by immediate and delayed recall, that much difference of opinion regarding correlation has arisen. A goodly number of researches have shown that much depends upon the method of testing recall; by choosing the method, one predetermines the result.**

In general these statements seem justified:

1. For this particular purpose, delayed recall should not be measured by relearning since the quicker learners, other things being equal, would excel, on that account, in the relearning test as well as in the original test.

2. The method of allowing the subject to complete the original learning by themselves—to report when they have "just learned it"—is faulty, since many really overlearn in various degrees the whole or, more often, certain portions. The test then, other things being equal, would show large differences depending on the amount of overlearning.

3. The method of giving a lesson which may be studied a given time, and by a standardized method, resulting in different amounts in immediate recall, seems preferable.††

*E. Meumann. *Psychology of Learning.* 1913 trans. by J. W. Baird. pp. 169 ff.

†*Educational Psychology,* vol. III. p. 362.

**The most complete account of work on this question will be found in Lyons, D. O. "*The Relation of Quickness of Learning to Retention,*" Archives of Psychology, No. 30, 1916.

††See Lyon, *op. cit.* p. 16 ff.

Elementary school children (grades 3–8) served as subjects in the present study. The material was of two sorts, non-sense syllables and connected sense material in the form of biographies. The children worked in squads of about 8 each, the amount of material at each lesson being such that the best subjects could recall from 75–90% immediately. The tests with non-sense material were repeated by each subject on 5 different days; in the case of sense material 6 tests were given. For the series of tests with each material, a revolving schedule was employed which neutralized the effects of unequal difficulty of texts, fatigue, diurnal differences and the like. The measure for each individual was the sum of the scores from the 5 or 6 tests. Details with regard to materials and methods will be found elsewhere.*

The following table summarizes the schedules:

Grade	No. of Inds.	Material	No. of tests	Length of each test	No. of hours before retention test
8	43	non-sense	5	9′	4
8	43	sense	6	9′	4
6	39	non-sense	5	9′	3
6	39	sense	6	9′	3
5	38	sense	6	9′	3
4	37	non-sense	5	9′	3
4	37	sense	6	9′	4
3	42	sense	6	7′30	3

The time of the study periods was kept constant (9 minutes),† the score being determined by the amount immediately recalled. Methods of study were defined, viz.: all must employ the "whole" method, first reading without attempted recall until, following a signal recall was employed as far as possible. Each student studied at his own tempo, and with or without inner speech as he chose. Retention was tested 3 or 4 hours later (always the same for a given class, of course) by simple recall in the case of nonsense syllables and by writing on the board names of the individuals whose biographies were studied to serve as a guide for recall in the case of the sense material.

***Recitation as a Factor in Memorizing.** Archives of Psychology, No. 40, 1917.

†Except for Grade 3 in learning sense material in which the period was 7′ 30″.

The coefficients of correlation were computed by means of the formula for rank differences:

$$\rho = 1 - \frac{6 \Sigma D^2}{n(n^2-1)}$$

From the values of ρ, the Pearson coefficient r was computed by

$$r = 2 \sin. \left(\frac{\pi}{6}\rho\right)$$

The coefficients of correlation between amount recalled immediately after the learning and amount recalled 3 to 4 hours later, are as follows:

Non-sense Syllables

8th Grade — +0.73 ±P. E. 0.04
6th Grad — +0.78 ±P. E. 0.04
4th Grade — +0.73 ±P. E. 0.05
Average +0.7 ±P. E. 0.043

Sense Material

8th Grade — +0.81 ±0.04
6th Grade — +0.86 ±0.03
5th Grade — +0.79 ±0.04
4th Grade — +0.89 ±0.03
3rd Grade — +0.80 ±0.04

Average +0.82 ±0.036

The evidence is clear that a substantially positive correlation exists between amount recalled immediately and amount recalled 3 or 4 hours later. This has an interesting bearing upon practical affairs but, of course, it is by no means a solution of the relation between "quickness of learning" and "retentiveness." Native retentiveness and other factors being equal, those who recall more immediately after study would recall more after an interval.

Next, consider the correlation between amount recalled immediately and the *proportion of that amount* recalled after the interval. The coefficients of these correlations follow:

Non-sense Material

8th Grade—	+0.41+P. E.	0.09
6th Grade—	+0.35+P. E.	0.09
4th Grade—	+0.42+P. E.	0.09
Average	+0.39+P. E.	0.09

Sense Material

8th Grade—	+0.46+P. E.	0.08
6th Grade—	+0.42+P. E.	0.09
5th Grade—	+0.28+P. E.	0.10
4th Grade—	+0.51+P. E	0.08
3rd Grade—	+0.40+P. E.	0.09
Average	+0.41	0.09

While these coefficients are not as high as those obtained between amount recalled immediately and amount recalled after 3 or 4 hours, they are high enough and consistent enough to indicate a positive relation. It is my belief, after examining the original data in detail, that one factor has some potency in reducing the present coefficients toward zero. Some individuals *overlearn* certain portions of the material unnecessarily before learning other portions. This would, of course, reduce the total amount immediately recalled and thus move these subjects toward the lower end of the scale. The fact of overlearning, however, would, as regards *percentage recalled later* move them toward the upper limit in that respect, not because they possessed superior retentiveness in general, but because they *overlearned a little rather than just learned much.* The subjects were requested, of course, not to do this, and were required to study by the whole method, but there is certainty that several did overlearn considerably. This seems to be indicated by the fact that the coefficients (following section) of the *percentage* recalled after 3 or 4 hours with teachers' estimates of ability are less than the coefficients of *amount* immediately recalled or the *amount* later recalled with the teacher's estimates. Moreover, these coefficients are subject to another source of attenuation not found in the coefficient between amount immediately recalled and amount recalled later. Neither measure of recall is absolutely accurate, each being subject to various variable errors. When *amounts* are

correlated, the coefficient is subject to attenuation according to the unreliability of each individual's score in each of the two tests, *i.e.*, immediate test and delayed test. But the measure of proportion between the two is subject to the combined errors of both since it is a statement of relation between them. This means an increased attenuation of the coefficients between amount immediately recalled and the proportion of that amount later recalled.

Taking into account the two sources of additional attenuation found in the correlation between amount immediately recalled and the proportion of this amount later recalled, it appears that these coefficients fall considerably below the truth. I should estimate that the coefficients as given above should be raised by a fourth, giving an average coefficient for each kind of material of about +0.50.

Correlation of Immediate and Delayed Recall with Teachers' Estimates of General Intelligence

Teachers of the various grades tested furnished estimates of the general intelligence of their pupils. They used their own conception of "general intelligence" but were urged not to identify it with marks attained in class work and to disregard the pupil's age. The estimate of but one teacher was obtained for each group.

COEFFICIENT OF CORRELATION OF TEACHERS ESTIMATES WITH

	Immediate recall non-sense	Immediate recall sense	Delayed recall non-sense	Delayed recall sense	Percentage retained non-sense	Percentage retained sense
8th Grade.........	+0.55 **P. E. 0.07**	+0.55 **P. E. 0.07**	+0.44 **P. E. 0.08**	+0.50 **P. E. 0.08**	+0.43 **P. E. 0.09**	+0.45 **P. E. 0.09**
6th Grade.........	+0.60 **P. E. 0.07**	+0.54 **P. E. 0.08**	+0.57 **P. E. 0.07**	+0.56 **P. E. 0.07**	+0.40 **P. E. 0.09**	+0.50 **P. E. 0.08**
5th Grade		+0.53 **P. E. 0.08**		+0.59 **P. E. 0.07**		+0.48 **P. E. 0.08**
4th Grade.........	+0.14 **P. E. 0.10**	+0.28 **P. E. 0.09**	+0.30 **P. E. 0.09**	+0.19 **P. E. 0.10**	+0.21 **P. E. 0.10**	+0.28 **P. E. 0.09**
3rd Grade.........		+0.48 **P. E. 0.08**		+0.50 **P. E. 0.08**		+0.48 **P. E 0.08**
Average..........	+0.43	+0.48	+0.44	+0.47	+0.35	+0.44

Correlations with teachers' estimates of general intelligence vary from +0.14 to +0.60, two-thirds of the coefficients falling between +0.40 and +0.60. The low values are nearly all produced by the 4th grade. For the other grades the central tendency is approximately +0.50. The different measures of recall correlate about

equally with estimated intelligence, the apparently low coefficients with *percentage recalled* being due to the greater attenuation which obtains in this measure, as just explained.

The findings here obtained are not in conflict with the results of most earlier studies, although some very different values varying from almost perfect to almost no correlations may be found. The median results are most numerous and support a prevalent opinion that (as the present findings indicate) immediate and delayed recall are positively but not perfectly correlated with general intelligence.

The following coefficients indicate the relation of results from the two forms of material used, *i.e.* unconnected non-sense and connected sense material.

	Immediate recall non-sense	Delayed recall non-sense
Immediate recall sense	8th $= +0.57 \pm$ *P. E. 0.07*	8th $= +0.58 \pm$ *P. E. 0.07*
	6th $= +0.49 \pm$ *P. E. 0.08*	6th $= +0.48 \pm$ *P. E. 0.08*
	4th $= +0.47 \pm$ *P. E. 0.08*	4th $= +0.24 \pm$ *P. E. 0.09*
Delayed recall sense	8th $= +0.50 \pm$ *P. E. 0.08*	8th $= +0.54 \pm$ *P. E. 0.07*
	6th $= +0.46 \pm$ *P. E. 0.08*	6th $= +0.58 \pm$ *P. E. 0.07*
	4th $= +0.43 \pm$ *P. E. 0.09*	4th $= +0.25 \pm$ *P. E. 0.10*

A positive correlation, large enough to bear some significance, exists between immediate recall of sense material and immediate recall of nonsense material and between delayed recall for the two materials, and likewise between immediate recall of one kind and delayed recall for the other kind of material. Psychologists, I venture to say, would expect a positive but far from perfect correlation (such as here shown) between such functions. These coefficients are about the same as those between any of the tests and a teacher's estimate of general intelligence or ability.

Summary of Results

(1) A postive and high (0.73 to 0.89) correlation exists between immediate recall and delayed recall when both functions are measured by absolute amounts.

(2) A positive correlation of at least 0.50 exists between the two functions when immediate recall is measured in terms of absolute amount and delayed recall is given in terms of the proportional relation between the two.

(3) All tests of immediate or delayed recall show a correlation with teachers' estimate of general intelligence with a central tendency of about +0.50.

(4) Coefficients of approximately +0.50 are found between the various single tests for the two kinds of material, senseful and nonsense.

(5) The results so far as they go, give a clear confirmation of the contention that correlation rather than compensation is the rule. Meumann's belief that "the typically rapid learner is usually a rapid forgetter" and that "the typical differences persist no matter whether the material to be learned is of a significant or non-significant sort" finds no support in the data.

PEDAGOGICAL SUGGESTIONS FROM MEMORY TESTS

T. L. PATTERSON

Department of Physiology, Queen's University, Kingston, Canada

A complete act of memory requires that impressions shall be retained, recalled, and recognized as familiar and as belonging with certain other impressions. The perfectness of any act of memory depends upon the kind and intensity of the impressions and of the associations between them.

Impressions are of six kinds, namely,—visual, auditory, motor, tactual, gustatory and olfactory, but the first three named are of the most importance in memory. Certain experiments have been made to determine which of these three kinds of impressions are best retained, and to discover the relation existing between recall and recognition.

One of the methods consisted of selecting thirty names of common objects and arranging them in three columns of ten each, care being taken not to put together words commonly associated. The test was made upon the pupils of a typical school and college in all grades from the third primary up. The words in the first column were pronounced to them at the rate of about one every two seconds: those of the second column having been previously writen upon the board, were uncovered one at a time at about the same rate and rubbed out as soon as the next was uncovered; while the objects named in the third column were shown at the same rate. In each case after the ten words were given, the pupils wrote as many of them as they could remember. Three days la er and at the same hour, the pupils were asked to write as many words as they could remember from three columns of ten wo ds each, which were similar to the three preceding columns and which were given orally. In the first column they were asked to think of the sound; in the second the visual appearance suggested; and in the third the objects named.

The following were the averages for the first three lists of words, namely,—6.85, 6.92 and 8.28 respectively, showing that objects were remembered better than the written names, nd the latter better than the spoken names. The averages for the last three lists of words were, 6.98, 7.91 and 7.48 respectively, showing that the visual qualities are remembered better than sounds and also better

than objects imaged. In reproduction of the same after three days it was found that mental images of objects were remembered better than their names, which is of great pedagogical significance, indicating that if objects are shown children, or when that is impracticable, if they are led to form mental images of them, they can obtain a genuine knowledge of things more readily than they can be crammed with the verbal appearance of knowledge. There is an increase from the primary department up, in the power of immediate reproduction of words of all kinds but the differences of only about two words between primary pupils and college students is not very great when we remember that it is not so much a matter of memory as it is of mental grasp, and that the younger pupils require longer time to write the words and hence would be more likely to let some of the words drop out of consciousness. In reproduction after three days the college students show no superiority over the children, and perhaps this may be due to the fact that the words were more deeply impressed upon the younger pupils than the older.

Another memory test along on these lines employed series of numbers instead of words. During the Spring of 1891, by permission of the School Board in Worcester, Massachusetts, Dr. Franz Boas of Clark University took certain anthropological measurements of the pupils in the grammer schools, and also made certain tests of eye-sight, hearing and memory. The memory tests which were made upon about fifteen hundred pupils in the grammar schools, above the second grade and below the high schools, together with some tests from the normal school, came into the hands of Professor Bolton for xamination. In addition to this and in order to complete the material for all the grades in the public schools, the tests were made upon some of the senior and sophomore pupils in the High School.

The method of making the tests was as follows: A series of numbers in which the digits were so arranged that they did not stand in their accustomed order and no digit was repeated, was read before each class to be tested, and each class was tested on four different occasions. The tests were made in the morning with the exception of two grammar schools, and the normal school, where the purpose was to determine the effect of fatigue, and here two tests were taken in the morning immediately after school assembled and the other two just before the closing of the session in the afternoon. The digits were dictated slowly and distinctly at intervals of about

two-fifths of a second with care to avoid rhythm or grouping, and at a given signal after the dictation of each number was finished, the pupils wrote the digits as they remembered them. In the lower grades of the grammar schools and in the high school twelve observations (a group of five or more digits) constituted the test, and in the eighth and ninth grades only nine observations were made at each test. In the lower grades the first three observations of each test were made with five digits; the second three with six; the third three with seven; and the last three with eight, while in the higher grades and in the high school the first three observations were made with six digits, and in the normal school with seven. The tests being repeated at four different times twelve observations with five, six, seven and eight digits respectively, were made upon each pupil and it would be possible for each pupil to get any number of the twelve correct. An observation was considered correct when only those figures which the teachers had dictated were present in the same order as that in which they had been dictated. It was found that when the observations at each test were begun with five place numbers and gradually increased to eight, the pupils easily grasped the five-place numbers and were led by these to grasp and retain more than they would otherwise be able to do. If the observations were begun with the seven-place numbers before making one with an eight or a nine, the number of figures was not so easily remembered, and more errors resulted from this cause. By classifying the pupils of a particular age or in each grade according to the proportion of the twelve observations on five-place numbers that were correct, and the same for six, seven and eight-place numbers, and marking the percentage that each class was of the total number of pupils in the grade or of the age on thirteen ordinates (twelve for the twelve obsevations and one for those pupils who had none correct) and connecting these points by a line, a curve representing the distribution of the pupils of the age or grade could be obtained, and the maximum of the curve would then show the proficiency of the pupils for the age or grade in remembering five, six, seven or eight-place numbers, as the case might be. Therefore, the various curves which were obtained showed not only the distribution of the pupils according to the ages and the distribution according to grades, but the increasing accuracy with which the older pupils remember a given number of digits. Considering that the tests measure the length of the memory-span, we can conclude that the memory-span increases with the age rather than with the growth

of intelligence, as determined by the tests used in promoting pupils from one grade to another. These tests however do not apply to the retentiveness of the memory, but may be considered more as tests of the power of concentrated and sustained attention. Observations upon the pupils while the tests were being made, seem to indicate that most pupils depend upon their powers of visualization to remember the number of digits, and at the same time they were noticed to repeat the digits as they were dictated.

A comparison of the standing of the pupils in their grades and their ability to remember figures was undertaken to detemine what was the relation between the memory-span and intellectual acuteness of the pupils. As the pupils depend upon their power of visualization, this subject becomes more interesting in determining how far this power is of service in school-work and how closely the power of concentrated and sustained attention is related to intellectual acuteness. For this purpose the teachers in the Oxford and Freeland Street Schools and of the High Schools of Worcester, Massachusetts, were requested to give an estimation of what they considered was the general standing of their pupils with respect to the school work. The pupils were classed as either good, fair or poor, and these classes were compared with three classes which were determined in the memory test by the proportion of correct judgments. The letters of the following table (A representing good, B fair, C poor,) in the left hand column indicate the teachers' classification, and those in the upper line the classification by the memory test.

Table giving the comparison of the teachers' estimation of standing of the pupils, and their standing, as determined by the memory tests. The classes of the teachers are represented as 100 and the others are expressed in percentages.

	A	B	C
A	32.6%	51.0%	16.3%
B	21.4%	58.2%	20.4%
C	24.1%	49.4%	26.5%

Only eight and a half per cent. more of the pupils who were classed A by the teachers have been classed A rather than C by the memory tests. If these tests may be considered as tests of the ability for concentrated and sustained attention and of the power of visualization, we can conclude that these powers are not the only ones concerned in intellectual pursuits and are not sufficient for a successful understanding of intellectual work. Intellectual acute-

ness, while more often connected with good powers of visualization and of concentrated attention, does not necessarily require them, and it cannot be said that those pupils who are bright intellectually are more distinguished on account of their good memories. The fact that a good memory is not necessarily accompanied by intellectual acuteness, adds weight to the conclusion that the growth of the memory does not necessarily accompany intellectual advancement.

In psycho-physical experiments it is customary to take seventy-five per cent. of right answers as the point at which the subject may be safely said to have some knowledge of that concerning which he judges. This standard is chosen for experiments in which a choice is made between two alternatives, where, by mere guesses, the subject will get 50 per cent. correct, but in this test the subject must be supposed to have exact knowledge before he can recall correctly any number of digits. Whatever standard is chosen for these tests, it must be considered as the probability that a certain number of digits should be judged correctly every time. If we then choose seventy-five—though it seems a less figure might be chosen—the results show that all pupils below the 6th grade and over thirteen years of age reach the limit of their memory span at six, and all others at seven. Six may be taken as the limit to the memory span for most grammar and high school pupils.

In order to treat this subject properly a comparison must be made between the boys and girls. For this purpose the boys and girls were classified according to their ages, and in order to get classes sufficiently large to form a comparison it was necessary to put the pupils differing by two years in age instead of one into each class. The probabilities of error were found for each class, and the comparison is made in the following table. Also the ages together with the probabilities of correct judgments are given for each class.

Table showing separately the probability of error for boys and girls.

Five-Place Numbers

	Boys	Girls
Pupils under 10 years	16.	13.7
Pupils over 10 and under 12	10.4	11.5
Pupils over 12	11.5	11.6

	Boys	Girls
Six-Place Numbers		
Pupils under 11 years	47.6	47.9
Pupils over 11 and under 13	35.2	26.5
Pupils over 13	25.4	25.
Seven-Place Numbers		
Pupils under 12 years	65.6	61.6
Pupils over 12 and under 14	50.7	51.
Pupils under 14	51.	44.
Pupils under 14 years	82.4	64.7
Pupils over 14 years	65.4	65.8

From this table it will be seen that in a majority of classes the girls made a decidedly less error than the boys, and in the classes where the boys surpass the girls, it is by a very small figure. This conclusion also harmonizes with the results of other observers.

Tests were also made to determine the unconscious memory and the effect of fatigue in three grammar schools, namely,—Oxford, Freeland and Woodland Street Schools. In the Oxford Street School four tests were taken in the morning and a different series of digit-groups are used at each test. The same digits dictated at the first test were read in the inverse order in the second. They were then completely re-arranged for the third and read in the inverse order for the fourth. Thus the digits in every observation were the same for the four tests, the order alone being changed. The same scheme was used in the Freeland Street School, only two tests were taken in the morning immediately after the school assembled, and two just before closing in the afternoon. In the Woodland Street School the same digit groups were used for the four tests, the purpose being to determine the effect of unconscious memory. The results from all the schools point to the conclusion that the pupils improve with practice. The great uniformity and large increases with each test in the Woodland Street School seem to show that the pupils unconsciously remember the digits that have been dictated one day previous. The probabilities of correct judgment do not show any variations due to fatigue.

The total number of correct judgments for the morning tests in the Freeland Street School was 2,690 and for the afternoon tests 2,640; for the morning tests in the Woodland Street School 6,609, and for the afternoon tests, 7,179. When we consider that great increases were made with each test, and the first test in the Freeland Street School was made in the afternoon, we should expect a greater number of correct judgments for the morning test; and

since the first test was made in the Woodland Street School in the morning, we should expect a greater number of correct judgments for the afternoon tests. This is just what the figures showed, and it may be safely concluded that the pupils suffered no fatigue from their school work, at least none discoverable by such tests as these. Their work was probably not excessive.

A careful examination of the observations show that the types of errors are three and perhaps four in number and these classes of errors represent stages in the fading of the memory-image. In the first place the digits suffer a displacement of order; in the second, other digits are substituted for some that were dictated, and in the third the number of digits is misjudged, either over or under-estimated. Various causes may be assigned for the displacement of order, as when a pupil attempts to write, the attention passes over the successive digits in memory as a rule much faster than they can be written. Before the pupil can write the first digit the attention has passed to the third or fourth and the hand is innervated for the digit that is present in consciousness. The second may be immediately recalled and put in the third place. However, it more frequently happens that the fourth or fifth is displaced than the second or third. Again, the order of the digits in the numbers previously dictated clings in the mind and causes the figures in the next number to be interchanged in accordance with the order. To illustrate,—Two numbers, the first commencing with 8163 and the second with 5136, were dictated. The 3 and 6 in the second were frequently reversed so as to read 63. Further, the order in which the digits stand in our system of notation determines some changes, as for instance, the last three digits of one number were 768 and a very frequent error made was to change the order of the 7 and 6 to 67. A fact that has been noticed frequently in teaching children and also adults is the great liability to confusion, when it is attempted to keep separate two like organs whose functions are diametrically opposed. As in physiology it is a difficult matter for children to distinguish the functions of the right and left ventricles of the heart and even for adults the functions of the dorsal and ventral columns of the spinal cord. It seems probable that this difficulty also appears in keeping the order of two digits that are easily remembered. The inversion of the order is by far the most frequent error as it is also the first to occur. In the second case that of substituting a new digit for one that has been read, there enter some of the causes that bring about an inversion of the order. A digit is substituted for another to make the two stand in the order

they have in the system of notation, or in the order in which they were in the number previously dictated. The likeness in the sounds of the names of two digits often determines the substitution of the one for the other, as 9 and 5 and 9 and 1 are frequently interchanged. The written or printed forms of 9 and 1 probably have something to do with the substitution of the one for the other. The frequent interchange of 3 for 8 is undoubtedly due to the likeness in the form of the printed digits. When a pupil has forgotten the proper digits and also their associations and so drops them out altogether, it shows a more advanced stage in the disappearance of the memory-image. Thirdly, when the pupils over-estimate the number of digits, two tendencies only were noticed. The digits that were supplied were put in the places in the notation system that occurred between two digits already given, or they were placed between two digits which should stand together and which were separated by the supplied digits in some number previously dictated, and when two digits already stand in their natural order, the tendency is very strong to put another digit in order either before or after those given. The second tendency was to repeat some digit already written. Over-estimation of a number was very infrequent, probably for the reason that each test was begun with numbers that could be easily grasped and digits that could be counted.

In the normal school the observations were begun with seven-place numbers; but instead of making three observations, as was done in the grammar schools, with seven digits, the teacher dictated only one seven-place number before dictating an eight and a nine-place number. Again, a seven- and an eight- and a nine-place number were dictated and so on until fifteen observations were made at each test. As results have shown that six figures are all that the best pupils can easily span, the normal school pupils were taxed to the limit of their powers on the first trial. In the grammar schools the pupils were started with numbers they could easily grasp and were led by steps to expect the number of digits in each succeeding observation. On this account 180 pupils from the Oxford Street School over-estimated the number of digits on observation with seven and eight-place numbers 88 times; and 24 normal school pupils over-estimated the number of digits 76 times. This tendency, however, to over-estimate shown by the normal school pupils is not the general rule, as the experiments of other observers like Drs. Hall and Jastrow, have shown that the tendency was to under-estimate the number of clicks made by a quill held against the notched circumferences of a revolving wheel, when the

number of clicks was too rapid to be counted. It would seem then from the preceeding experiments that at the moment when pupils reached the limit of their memory-span, they would over-estimate the number of digits; but if the experiments had been continued with a greater number still, the pupils would have under-estimated the number.

To determine the class of errors which was first to occur, all those observations in which only one kind of error was made, were examined on 105 tests from the fourth and fifth grades.

Table showing the different classes of errors,

	Inversions of orders	Substitution	Over-estimation	Under-estimation
Five-place Numbers	10	7	3	1
Six-place Numbers	44	15	15	5
Seven-place Numbers	40	26	13	2
Eight-place Numbers	33	10	14	4

If we take observations with a much greater number of digits we should expect that the greatest number of errors would result from the dropping of digits The total number of errors of every class in the Oxford Street School, when compared with the total number of observations, give the following percentages: Inversion of the order, 53 per cent.; substitution, 31 per cent.; under and over-estimations of the number of digits, 15 per cent.

It is very difficult to tell at what places errors are most likely to occur, or in what places the digits are most often forgotten, but as the well-known rhetorical principle runs that the emphatic words in a sentence are the first and the last, so it is with numbers; the first and the last are emphatic and those figures in between are those that are forgotten.

Another method for testing the memory of children at different periods of school life, and to determine what appeals to the different senses and to the sympathies of different ages was that of a story, designed to appeal about equally to different senses and faculties, which was written by Dr. G. Stanley Hall, and divided for examination as here indicated: James: Mack: ten years old: a farmer's son: dreamed: that his father: and mother: died: very poor: and left him nothing: but 37 cents: a loaf of bread: and a Bible: etc., etc. The story contained 324 words and was divided into 152 parts, which were called "terms" for convenience. In this division of the story it was sought to have as many terms as there were distinct facts or ideas. About 700 tests were made in the following schools of Worcester, namely: Woodland Street, Downing Street,

Oxford Street, Winslow Street, and Salisbury Street grammar schools; both high schools, English and Classical, and in one department of Clark University, together with Miss Aiken's School at Stamford, Connecticut. Effort was made to avoid disturbance and to have the same conditions for all; and the pupils were encouraged to make the best possible effort. When a room was entered preparatory to making the test, the pupils were told that they were to be given a memory test, in which a story would be read, which would take three minutes. They were requested to give the best possible attention and told that the same test had been, or would be, made in other grades and in other schools. The pupils were requested neither to give nor receive aid, to write on sheets of paper previously placed on their desks as soon as the reading was finished, and to hand the papers to their teacher as soon as they had written all they remembered of the story. In the seventh grade an additional test was made in which the pupils read for themselves. It was made the same hour in the day and, as far as possible, under the same conditions as those where the story was read to them. Typewritten copies of the story were placed on each desk, with the printed side down, together with blank sheets of paper. At a given signal the pupils turned the papers over and were given three minutes to read them. This was sufficient time to read the story at a slow rate and many commenced a second reading, and some probably read the entire story twice in this time. At a second signal they turned the papers back again and commenced writing. The test was not merely one of word memory, nor was it a test in reproducing the thread of the story. A rigid adherence to either plan would in part have defeated the end sought. In the first case would have been the difficulty that many synonyms would have been used to express the same idea, which could not be said to be lost from memory. In the second case, if strict adherence to the order of the story had been required, many papers would have been thrown out, changing the entire nature of the results. All terms which fell below this standard were not counted; but an attempt was made to maintain, as nearly as possible, the same standard in the study of all the papers. Sometimes a sentence would be so constructed that it would contain only one term that could be counted. In such a case the term was not counted, unless it retained its original meaning and it was of such a nature as to indicate that it had its origin in the story. For example, "James dreamed that his parents died," "Parents" was credited for both father and mother. In descending from the higher to the lower grades there was a decided

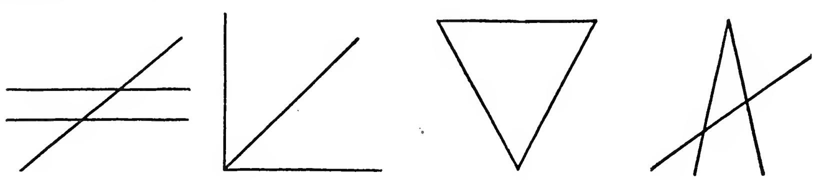
tendency to increase the number of words used to express the same number of ideas. The test was only partially successful in determining what was expected. The appeal to the sympathies and senses was not sufficient to warrant any general conclusion in this direction. No difference between boys and girls in this respect was found, but other important results were obtained. One of the most striking facts brought out was the early age at which children reach approximately their maximum memory power. The boys in the second year of the high school remembered 43 per cent., the highest per cent. for boys. The next highest was 42 per cent. in the ninth grade, while the seventh grade made 37 per cent., one per cent. less than that for the fourth year of the high school. Under these conditions it seems that boys reach their maximum memory power early in the high school period. The girls remembered 43 per cent. in the seventh grade and each high school grade remembered 47 per cent. As a whole, so far as the grades ran parallel, the girls remembered 4 per cent. more of the story than the boys, and no decline of memory power was discovered in the girls. 37 per cent. of the whole story was the average remembered by all grades. This may seem a very low per cent., but it must be kept in mind that the story is a very hard one to memorize.

The highest average attained by any grade was 52 per cent., the amount remembered by a class of 30, in Miss Aiken's school, of the average age of sixteen. This high per cent. is perhaps due to special memory training given to the class. The girls on an average had been in the school for two years. The first third of the story was remembered better than the second third by all grades with one exception, and the second third was remembered better than the last third without exception. The relative difference for the three parts is much greater in the lower grades than in the higher ones. The lowest grade remembered about one-third as much of the last part of the story as of the first, while the fourth year high school remembered 70 per cent. as much of the last third as they did of the first. The marked difference in the seventh grade, where the pupils read for themselves, is probably due to their reading the first part twice. The per cent. remembered by all in the three-part division was successively 46, 38 and 27. Practically the same ratio was sustained by all grades. It seems from the results, that the memories of some pupils act spasmodically. They remember a portion with a considerable degree of accuracy; then they pass on for some distance without getting anything. Others touch, as it were, only the "high places" and so continue through the whole story. Some

with a limited power of reproduction bring out the least expected points and a considerable number remember the first part of the story quite well and remember very little in the latter part. Here fatigue seemed quite clear. Apart from the errors in spelling, in understanding the words, etc., the errors may be classed as errors of omission, insertion, substitution and misplacement. The parts left out were clear omissions, the insertions were due to imagination, the substitutions occurred most frequently where words used in the story had approximate synonyms in the child's vocabulary, and displacements occurred most frequently in the order of the personages mentioned, and where modifiers in one instance seemed well adapted in another. All degrees of inaccuracy were found in the paper, from mere failure to understand a word, to pure products of imagination. For example, James Mack became John Mack, Tommy Mack and Henry Mack, etc. For a story that was written for the test and which did not aim at literary form, but a difficult story for memorizing, the maximum memory power is reached at a relatively early age. The boys in the third grade remembered only 17 per cent. of the story. In the ninth grade they remembered 42 per cent. and in the high school about 40 per cent. From this it seems that memory power for the boys culminates about the beginning of the high school period. The girls made a rapid increase from 18 per cent. in the third grade to 43 per cent. in the seventh grade, and 47 per cent. in the high school. The class in Miss Aiken's school remembered 52 per cent., but this high average is probably due to special training, not usually given in public schools.

Another method of testing the memory was that used by M. W. Meyerhardt of Clark University during the year 1908, who tested respectively the Grapho-motor memory and the Visual-motor memory of college and university students. The method consisted in having the subjects count off numbers in rotation, each individual keeping his number in mind, and placing it upon a booklet with his name and age when passed to him. The booklet contained four figures, one on each leaf, and the subject was told not to open the book until the signal was given to do so. The object of the experiment was then explained and the odd numbered subjects were instructed to write the odd numbered characters and look at the even numbered characters; the even numbered subjects were to look at the odd numbered and write the even numbered characters during the time lapsing between signals which was 10 seconds.

After disposing of all four characters the books were closed and at the next signal (which was to follow 40 seconds after the last one) all subjects were to write the four characters once as they remembered them, on a sheet of paper which was provided for and given out with the book with their name, age and number to be placed thereon the same as on the booklets. The results that Meyerhardt obtained show that in nearly all cases the grapho-motor memory is stronger than the visual-motor memory and that college students are more particular with their drawings than university students. In order to carry the test one step farther the present author was enabled through the courtesy of Mr. G. F. Williams, Principal of Anson Academy, North Anson, Maine, to make about 50 such memory tests on his students in 1909 in the hope that from the results obtained he might further demonstrate the above and prove that not only the grapho-motor memory is stronger than the visual-motor memory, but that the high school students are as much more particular with their drawings, as the college students are more than the university students. The method employed was the same as that used by Meyerhardt with one exception, namely,—that a different set of test figures was used which nevertheless were of about equal difficulty for the subjects to reproduce. The figures used in this set of experiments are reproduced below.



1. 2. 3. 4.

From the results of the experiments on these students of high school grade, whose ages ranged from thirteen to nineteen years, it was found necessary to discard the results of 16 subjects, since they must have misunderstood the directions as given, or else interest was lacking. Probably it was due to the former, for they did not make any attempt at reproducing the figures in the booklet, and for this reason their results could not be counted in the final summary, since it was purely a test of the visual-motor memory and not a test between the powers of the grapho-motor and the visual-motor memories, thus defeating the object of the experiment. How-

ever, the results from the 34 remaining subjects may be said to have met the object of the experiment, and from these it was found that the results harmonized with those of Meyerhardt in that the grapho-motor memory is much stronger than the visual-motor memory. The results from 29 subjects or 85.3 per cent. showed this increase in grapho-motor memory over the visual-motor memory, while the results from the other five subjects showed little difference in their grapho-motor and visual-motor memory capacities. This may be accounted for in part by the above subjects not reproducing the figures but once or perhaps only partially, for in all cases where there were several reproductions of the same specific figure, much greater increase was indicated in the direction of the grapho-motor memory.

Therefore, since the grapho-motor power exceeds the visual-motor power in the capacity of acquiring knowledge through memory, all educational subjects from the graded schools to the university, as far as possible and as far as time will permit, should be taught from the grapho-motor standpoint, as well as by the other special senses. This of course cannot be applied to all subjects alike in the curriculum of our graded college and university school systems, but there may be a most satisfactory application of it in certain educational subjects like the biological sciences and in varying degrees to other subjects in our school curricula. In other words, teachers must not forget to train the grapho-motor sense or the memory obtaining capacity through drawing, as well as always finding receptive stimuli to develop the other special senses, for only along these lines can the results of teaching approach that point of maximal efficiency.

In general confirmation of the results of other observers, the girls were found in these experiments, to make a decidedly less error than the boys. The drawings of the figures as compared with some of those of Meyerhardt from the college students seemed to be at least on a par, and I would say in general terms, that there was a slight improvement over the drawings of those made by the college students.

Additional experiments on students of different high schools would of course have been advisable, but since the results of this set of experiments harmonize so closely with those conducted on students of the college and university grades, the results may be said to be conclusive.

COMMUNICATIONS AND DISCUSSIONS

A CLASS EXPERIMENT WITH THE HILLEGAS SCALE

The following study was made in order to illustrate, for classes in educational psychology, the use of the Hillegas Scale.* Twenty-five short compositions were taken as the material to be graded. These had all been written by college freshmen, and were all reproductions of the story of the "Marble Statue." The judges who did the grading were forty-one women, Junior and Senior students in educational psychology. They had had no experience in grading such compositions.

Each student had a copy of the twenty-five essays, and she was directed to grade them first without the use of any objective scale, giving to each any score from zero to one hundred which she felt that the composition deserved. A week later the use of a scale was explained, and these students were then asked to grade the same set of compositions by means of the Hillegas Scale.

Table I gives the results of the grading of the forty-one judges. In the first column is the number of each composition. Columns two, three and four give results without the scale, and columns five, six and seven give results with the scale. In columns two and five are given the variations in range of the judgments on each composition. In columns three and six are the medians of the judgments, and in columns four and seven are the standard deviations of the judgments on each composition.

An examination of this table shows the following:

1. The range of judgments is absolutely very wide, both with the scale and without it, but

2. The range is less with the scale. The average reduction is 11.5 points.

3. The range of the medians is the same, 32 points, with the scale and without it. Therefore, though the scale reduces the range of judgments on each composition, it does not do so at the cost of reducing the difference between the best and the worst compositions.

4. The standard deviations are reduced in the cases of all except two of the compositions. The average S. D. without the scale being 11.2 and the average S. D. with the scale 8.9.

These points would seem to show that the tendency of the scale is to make judgments more stable.

Another way of inquiring into the effect of the scale is to compute the correlation coefficients for various pairs of judges. It seems fair to assume that if the scale tends to promote uniformity it will

*M. B. Hillegas. *A Scale for the Measurement of Quality in English Composition.* Teachers College Publications.

TABLE I

Judgments of Forty-one Students upon Twenty-five Compositions

Without Scale				With Scale		
I. Composition no.	II. Size of Range	III. Median	IV. S. D.	V. Size of Range	VI. Median	VII. S. D.
1	45	65	11.1	41	67	9.0
2	50	65	13.0	35	67	9.7
3	55	75	10.3	48	76	9.4
4	47	85	10.7	40	77	9.8
5	58	68	12.8	47	65	10.2
6	42	75	11.2	37	72	8.6
7	50	75	14.5	38	72	9.6
8	45	68	11.0	33	67	8.8
9	50	70	11.8	38	68	9.1
10	50	80	9.6	38	75	8.6
11	23	90	5.9	22	85	6.4
12	70	75	12.3	36	78	8.8
13	45	80	8.9	32	75	8.5
14	60	74	10.8	41	73	9.1
15	66	58	16.9	35	55	9.9
16	70	60	16.0	48	67	9.9
17	35	80	9.6	32	77	8.1
18	53	75	11.4	45	75	9.8
19	62	69	12.7	40	69	9.3
20	45	80	10.3	37	75	8.1
21	70	70	16.0	55	68	11.5
22	30	85	8.0	36	83	7.1
23	45	70	8.9	34	72	9.4
24	39	80	8.7	34	77	7.8
25	30	85	7.7	25	87	6.9
Average	49.4		11.2	37.8		8.9

tend to raise the correlations between judges. In order to test this point the intercorrelations between forty pairs of judges were computed. For each pair the correlations were figured between their judgments before using the scale, and then between their judgments with the scale. In Table II the numbers in the first column stand for the various pairs of judges. The second and third columns give the Pearson coefficients as between judgments without the scale and with the scale respectively.

TABLE II

Correlations Between Various Pairs of Judges

Pair No.	Without Scale r=	With Scale r=	Pair No.	Without Scale r=	With Scale r=
1	.11	.36	21	.51	.63
2	.25	.49	22	.51	.39
3	.25	.32	23	.52	.39
4	.26	.00	24	.52	.28
5	.29	.26	25	.54	.32
6	.29	.45	26	.54	.67
7	.32	.52	27	.56	.53
8	.35	.19	28	.57	.55
9	.36	.54	29	.57	.39
10	.38	.29	30	.58	.67
11	.41	.35	31	.59	.70
12	.41	.42	32	.59	.51
13	.41	.48	33	.60	.62
14	.43	.27	34	.61	.62
15	.46	.55	35	.62	.63
16	.46	.44	36	.63	.20
17	.47	.45	37	.66	.24
18	.47	.39	38	.66	.78
19	.48	.64	39	.75	.52
20	.50	.57	40	.86	.70
			Average	.48	.46

From this table it appears that the range of correlations without the scale is from .11 to .86, and with the scale is from .00 to .78. The average correlation is .48 without the scale and .46 with the scale. In order to compare the judgments of different groups the forty-one students were divided into two classes of twenty and twenty-one members. The average scores of the twenty-five compositions as judged by the first class were correlated with the average scores as judged by the second class. This was done for the class judgments without the scale and with the scale. By the Pearson formula $r = .87$ in both cases. By the formula $1 - \frac{6 \Sigma D^2}{n(n^2-1)}$ the figures were .92 without the scale and .87 with it.

At first thought these correlation results seem to be in opposition to the results as shown in the standard deviations. But I doubt whether there is a true contradiction here. It seems to me compatible to say that, whereas the scale does steady the judgment of a group with respect to its marking of a single composition, it does not alter the judgment, either of the individual or the group, with respect to the relative merits of a series of compositions.

KATE GORDON.

Carnegie Institute of Technology.

AN EXPERIMENTAL STUDY OF CHARACTER

Having become interested in the "Experimental Study of Self-Analysis and Estimate of Associates" by Professor Hollingworth,* I have attempted a partial repetition of his experiment, with a somewhat different set of traits. My results, though they do not confirm Hollingworth's in all particulars, in general tend toward similar conclusions.

Twenty-one girls, all living in the same house, were asked to pass judgment upon various traits of character as possessed by their associates and themselves. At intervals of one or two days each girl was given twenty-one slips of paper upon which were written the names of her associates, her own name being included in the list. She was asked to arrange them in order of merit according to the degree of possession of some one trait. This was done for the following traits of character: pleasing personality, beauty, refinement, neatness, enthusiasm, optimism, thoughtfulness of others, leadership, self-esteem, and snobbishness. The experiment was carried out in such a manner that the girls considered their arrangements anonymous, which supposition led them to be honest in the placing of themselves. From the results the arithmetical mean rank was computed for each girl in each trait.

* * * * * * * * *

Hollingworth found a constant tendency toward greater accuracy in judging their associates than in self-estimation. I find this true for the five traits—pleasing personality, enthusiasm, optimism, leadership, and thoughtfulness of others; but do not find it so for the other four—neatness, refinement, self-esteem, and snobbishness as shown in Table I.

TABLE I

Trait	A. D. of Assoc.	A. E. of Self	Trait	A. D. of Assoc.	A. E. of Self
Pleasing personality	3.8	5.6	Neatness	4.2	4.2
Enthusiasm	2.8	4.4	Refinement	3.9	3.6
Optimism	3.5	6.1	Self-esteem	3.9	3.6
Leadership	2.8	5.5	Snobbishness	4.2	3.3
Thoughtfulness of others	4.4	5.7	†		

A. D. =average deviation of associates.

A. E. =average error of self-estimate.

Table II shows a tendency to overestimate themselves in refinement, leadership, enthusiasm, optimism; but to underestimate in the case of thoughtfulness of others, pleasing personality, self-esteem

*"*Vocational Psychology;*" Chapter VII.

†Due to an error in the experiment, it was necessary to omit the self-estimates for beauty.

and snobbishness. Hollingworth found over-estimation in all desirable traits and underestimation in all undesirable traits.

* * * * * * * *

TABLE II

Trait	No. over estimating	No. under estimating	Trait	No. over estimating	No. under estimating
Refinement	12	9	Thoughtfulness of others	9	12
Leadership	11	8	Pleasing personality	7	13
Enthusiasm	10	8			
Optimism	10	9	Self-esteem	7	9
Neatness	9	9	Snobbishness	5	16

*

*Due to an error in method of the experiment, it was necessary to omit the self-estimates for beauty.

By correlating each girl's judgment of every individual in the group with the average, or combined, judgment of the group, a coefficient of judicial capacity was found for each girl. This was correlated with her degree of possession of the various traits to determine whether one who possessed a given trait to a marked degree was a better judge of that same trait in others than was one who was herself deficient in the trait. As shown in Table III, those who possessed the desirable trait of thoughtfulness seemed to be unusually good judges of that trait in others. On the other hand, those possessing snobbishness to a high degree were poor judges of that trait in other people. Hollingworth found a positive correlation of .33 in the case of snobbishness where this experiment shows a negative correlation of .51.

TABLE III

Trait	Cor. betw. judic. cap. and poss. of trait	Trait	Cor. betw. judic. cap. and poss. of trait
Thoughtfulness of others	.87	Pleasing Personality	−.18
Leadership	.45	Enthusiasm	−.01
Optimism	.31	Neatness	−.26
Refinement	.30	Self-esteem	−.26
Beauty	.29	Snobbishness	−.51

In Table IV the various traits are correlated with one another. Those judged as having a pleasing personality were also ranked high in leadership, optimism, enthusiasm, and beauty. Those who were considered good in leadership were ranked high in pleasing personality, enthusiasm, and optimism. Beauty correlated merely with pleasing personality. Snobbishness is correlated negatively or in very slight degree with everything except self-esteem. Those who possessed the greatest degree of snobbishness and self-esteem were considered least thoughtful of others.

TABLE IV

	Pleasing Personality	Beauty	Refinement	Neatness	Self-esteem	Snobbishness	Enthusiasm	Optimism	Leadership	Thoughtfulness of Others
Pleasing Personality..		.72	.66	.44	.32	—.30	.73	.77	.87	.28
Beauty.............	.72		.41	.31	.49	.34	.48	.38	.56	—.20
Refinement..........	.66	.41		.68	.16	—.35	.41	.51	.61	.37
Neatness............	.44	.31	.68		.14	—.30	.23	.39	.50	.22
Self-Esteem..........	.32	.49	.16	.14		.63	.49	.31	.52	—.52
Snobbishness.........	—.30	.34	—.35	—.30	.63		.10	—.15	.05	—.66
Enthusiasm..........	.73	.48	.41	.23	.49	.10		.74	.87	.10
Optimism...........	.77	.38	.51	.39	.31	—.15	.74		.81	.23
Leadership..........	.87	.56	.61	.50	.52	.05	.87	.81		.18
Thoughtfulness of others..............	.28	—.20	.37	.22	—.52	—.66	.10	.23	.18	

Although these results may have little significance or reliability if viewed alone, they are sufficiently suggestive to indicate the desirability of continuing such experiments.

MINNIE MAE ROBSON.

Psychological Laboratory of
Northwestern University.

SCORING THE MONROE SILENT READING TESTS

Any test of a given ability is of value only in so far as it affords a reliable index of the degree of the ability present. What might otherwise be a very good test is sometimes vitiated by faulty methods of scoring. Not long ago the writer gave to a group of classes the Monroe Silent Reading Tests and the Trabue Completion Test, B Scale. The correlation between the pupils' scores in the two tests was found to be surprisingly low. The results in six classes were as follows:

School	Grade	Correlation (Pearson formula)
L	7 A	.15
L	7 B	.18
P	7	.36
L	6 A	.10
L	6 B	.11
P	6	.27

There is supposed to be the closest correlation between the Trabue Tests and general ability of any tests yet devised. What, then, are we to think of the Monroe Silent Reading Tests? It is my impression that the tests themselves are quite satisfactory, but that the author's directions for scoring are seriously at fault. There are numerous instances where the child surely comprehends

TABLE I

Pupil	Trabue Completion	Monroe Reading Strict Scoring	Monroe Reading Liberal Scoring
1	7	8.2	15.0
2	8	19.3	19.3
3	8	23.0	23.0
4	10	20.9	27.1
5	10	23.0	35.3
6	10	20.9	20.9
7	11	18.2	18.2
8	11	20.3	28.8
9	12	6.4	9.1
10	12	23.6	32.9
11	12	16.6	33.6
12	12	16.5	16.5
13	12	5.2	16.6
14	12	19.3	19.3
15	12	8.6	17.2
16	12	19.3	19.3
17	12	27.6	27.6
18	12	20.3	24.1
19	12	18.2	18.2
20	13	13.8	13.8
21	13	12.1	12.1
22	14	16.5	24.7
23	14	17.2	19.3
24	14	13.9	16.0
25	14	26.8	50.8
26	14	25.5	34.0
27	14	12.2	12.2
28	14	19.7	30.9
29	15	20.3	20.3
30	15	16.1	19.3
31	15	19.3	19.3
32	16	17.1	25.0
33	16	23.0	23.0
34	17	23.0	30.9
35	18	24.4	28.5
36	18	17.2	29.5
Median	12	18.7	20.7

what is read, but does not follow the exact directions in recording the answer. According to the author's directions for scoring the test such responses must be entirely rejected. A more liberal basis of scoring would give the child credit for the test whenever the answer indicates that the substance of the passage has been fairly well comprehended. From grade 6B in School L above 36 papers were received. Of these 21 receive higher scores when graded by comprehension than when graded according to directions. The differences are seen in Table I.

As a result the median of the entire group of 36 papers is raised from 18.7 to 20.7, and the correlation with the Trabue Completion Test is changed from .11 to .63. Most of the differences in the scores are due to the following directions: "If a pupil is asked to underline a word, the word must be underlined and not have a circle drawn around it or a check mark placed after it, in order for the exercise to be counted correct. If a pupil is asked to draw a line around a word, the word must have a line drawn around it in order for the answer to be counted correct." In the rapid work necessitated by the test it frequently happens that the child's attention leaps ahead to the next task, and he becomes careless about following the exact directions in regard to marking. Yet from the word he checks it is obvious that he has understood the passage. A more liberal basis of scoring, therefore, does greater justice to the reading ability of the child.

ERNEST C. WITHAM.

Southington, Conn.

PERFORMANCE NORMS FOR THIRTEEN TESTS*

This bulletin is a second contribution on the subject of mental tests, the first having been published under the title of "Eleven Mental Tests Standardized." Twenty-four tests have thus far been described and have been used to supplement other psychometric measures. The bulletin is the work of Investigators Dr. Gertrude E. Hall and Miss Catherine E. Conway who examined 757 individuals by the tests given. In addition to this, 314 school children were examined in an attempt to establish norms. The tests are:

1. Knox Cube Imitation: Test designed by Dr. Howard Knox, Assistant Surgeon of the United States Public Health Service, Ellis Island, N. Y. Material: Four 1-inch cubes four inches apart fastened to a thin board to form a base. One cube for S. and one for E. Movements: Movements are from S.'s left to right.

*Eugenics and Social Welfare Bulletin No. VIII. New York State Board of Charities, 1917.

Line	Tap Cubes
1	1, 2, 3, 4.
2	1, 2, 3, 4, 2.
3	1, 2, 3, 4, 2, 3.
4	1, 3, 2, 4.
5	1, 3, 4, 2, 3, 1.
(a	1, 3, 4, 2.
5(b	1, 3, 4, 2, 3.
(c	1, 3, 4, 2, 3, 1.

Summary: (a) Success may be expected of normal seven-year-old children on the first trial. (Fourth line.)

(b) Success may be expected of normal thirteen-year-old children when more than one trial is allowed. (Fifth line.)

2. Three-Number cancellation: This test is similar to the cancellation test described in Bulletin 5 of the State Board of Charities, Bureau of Analysis and Investigation, except that three numbers instead of one are crossed out. The test was designed to test more prolonged attention and concentration than is called for in the one-number test. S. is required to hold in mind the three numbers to be crossed out, to cross them out and to work as fast as possible. Thus the genuine value of the test seems to be what Meumann calls "real capacity of observation," and what Chambers calls "quickness of perception."

Summary: Children from the age of seven years can perform this test when unlimited in time; but when the time is limited to three minutes with from two to eight errors, success may be expected of normal thirteen-year-old children.

3. Recall of Objects: The idea of this test was obtained from the Neurological Institute test of a similar nature, but the method of presentation and scoring is different. The primary object is to test visual memory by using objects. Because of the secondary factors that enter into the test, namely attention, time and ability to follow directions, it was difficult for slow, unobserving children, especially those of the border-line and defective types.

Summary: Normal nine-year-old children can remember seven objects with one misplaced. Normal thirteen-year-old children can remember eight objects with none misplaced.

4. Grouping of Objects: The ten small objects used in the Recall of Objects test are used for an association test. This test was first used by the Neurological Institute. After the former test the ten objects are uncovered and S. is asked to arrange them in pairs, any way he wishes so long as he can account for the grouping afterwards. When S. has finished grouping the objects in pairs he is asked to give reasons for his grouping in the following manner: "Tell me why you put the shoe and the hat together," and so on until each pairing is accounted for in some manner.

Summary: Normal children from the age of seven years may be expected to group the shoe and hat, the gun and hatchet, and the dog with either the chicken or horse. The grouping of the cup, car, chair and either the horse or chicken may vary at all ages. Normal thirteen-year-old children may be expected to account for their groupings. Low grade morons and imbeciles fail on this part of the test.

5. Peg Design: The Bureau of Analysis and Investigation offers the peg design as a simple but interesting learning test. Retentive memory is also tested by the reproduction of the figure half an hour after the first success.

Summary: Success may be expected of normal nine-year-old children. High grade morons may perform this test, but require more time to learn the design than subjects of corresponding ages.

6. Story Reproduction: A simple story was selected by the Bureau for this test, the object being to find at what age children can remember the essential points of the story and reproduce a connected narrative.

Summary: Normal thirteen-year-old children may be expected to give sufficient details to show that they have gained the idea of the story. Morons fail on this test.

7. Syllogisms: To find at what age abstract reasoning may be expected, the Bureau adopted five syllogisms as a test. It was not expected that young children could do much with them, but they were given to the young children just the same as to the older ones to establish their reaction.

Summary: Normal thirteen-year-old children may be expected to form the conclusions of three of the five syllogisms.

8. Four-Detail Drawing: In the belief that drawing is a natural means of expressing ideas, the Bureau of Analysis and Investigation devised the following drawing test, which it is hoped will be a useful supplement to those tests in which ideas are expressed verbally, by writing and by actually doing things. The story chosen has a situation involving action, plenty of human interest, and allows freedom of expression. It was selected from Book I of the Mother Tongue by Arnold and Kittredge, and is entitled "Kindness to Animals." In this test the story was read, a passage of the story was written on the blackboard and the children were then told to illustrate the passage by drawing and to use as much of the story as they could remember to help them make a good drawing. Success depended in the first instance upon the attention given, and upon auditory, visual and retentive memory.

(a) Children with mental ages below X years stand a poor chance of succeeding with a test of this sort.

(b) In the group of children who were tested mentally, it is not the "at age" children, that is the children whose mental ages correspond to their chronological ages, who form the larger per cent. of successes.

(c) Disregarding the over-age or retarded group,

Mental Age	Per Cent. Successful
XII	95
XI	79
X	57
IX	20
VIII	30
VII	8

Summary: Success may be expected of normal fourteen-year-old children.

9. Three-Detail Drawing Test: This test is similar to that described in No. 8, but is for younger children. A simple story is read which contains a specific statement of facts which are used as a basis of the three-detail drawing test. The whole story is read to the children, the special quotation containing the detail was written on the blackboard, and the children are told to draw the three principal details in the quotation. The papers were classified by grades to find out if a curve of successes would correspond to that in the four-detail test.

Summary: The average child found in the second grade, or normal eight-year-old children can perform this test.

10. Balance Nickel (This and the following tests were worked out by Investigator Jessie L. Herrick, M. D., in the Dunkirk Public Schools): The need of performance tests for young children caused the Bureau to experiment with balancing a nickel. A new coin balances too easily and a very old one that is worn off on the edge will not balance at all. Therefore, a moderately used coin is best to use, and if this test should become standard it would be necessary to have the coins to be used tested as to flatness of their edges. It is probable that children will take more pleasure in trying to balance a real nickel than a metal disk of like size. A level table is necessary and a stop watch is used for timing the reaction.

Summary: All normal six-year-old children can balance the nickel when two hands are used. Normal children from the age of seven years can do it with one hand.

11. Peg Board: The idea of using the kindergarten peg board as a motor co-ordination test was gained from the Neurological Institute psychologists. The peg board is of wood, six inches square, 3-8 of an inch thick, with a square of 100 holes bored in it, ten in a row, 1–2 inch apart, the holes being 1–8 of an inch in diameter and a quarter of an inch deep. One cylindrical wooden peg is used, 1–16 of an inch long and of a size to fit the hole. The end of the peg is whittled a little toward a point so that it will enter the hole without resistance, for the speed is materially lowered by friction in the insertion of the peg. For perfectly standard work a metal board and stylus made to exact proportions would be de-

sirable. However, an examiner becomes familiar with his own tools and can easily detect individual differences in the persons he examines. A stop watch is used to time the reaction.

Summary: Time and method are the determining factors in this test. All normal children from the age of six years can do it. Ten-year-old children can do it in 65 seconds using the alternative method.

12. Tower: The tower test is one planned by the Bureau to use with young children, and is in three sections, the building of the tower, raising it from the floor to the table, and packing it together in the smallest possible compass. A nest of boxes such as is sold at ten-cent stores is used. The outer box is 3¼ inches or 3½ inches square, and there are seven or eight boxes diminishing in size until the smallest is 2½ inches tall and slightly less than an inch square. Each box except the smallest has one side left open. The boxes are covered with brightly colored pictures which please the children. Moreover, the toy is of a kind familiar to most all children and so raises no fear in them.

Summary: Build: Normal ten-year-old children may be expected to build the tower in an average time of 28 seconds.

Raise: Time and the number of blocks off are the principal elements in this part of the test. A normal ten-year-old child may be expected to raise the tower in 58 seconds with no more than two blocks off.

Pack: A normal nine-year-old child may be expected to pack the blocks in an average time of 22 seconds.

13. Boat: Four pieces of cardboard, of a size that can be kept in a small envelope. A boat is marked 300 lbs., meaning that the boat will carry three hundred pounds, a large man, usually referred to as "the fat man," marked with his weight, 300 lbs., two little men, just of a size, each marked with their weight, 150 lbs.

Summary: A normal ten-year-old child or a subject with a ten-year-mentality may be expected to perform the boat test with not more than two protests.

With the excpetion of the last four tests the working norms were obtained from the results of a selected group of public school children, all of whom had made the average yearly progress from grade to grade. The results of the work of the institution group, for the most part do not come up to expectation when compared with that of the public school group, but this is due to the fact that in the former group the backward children tend to lower the averages of those children in the institutions whose work is comparable to the normal public school children. The results show that in tests involving reasoning and prolonged concentration, the institution children as a group make a poorer showing than the selected groups of public school children, while with simple performance tests and tests that are quite mechanical the results of the two groups are more nearly alike.

Mental Examinations—Bulletin No. XI.

This Bulletin concerns mental examinations of orphan asylum children. Psychometric examinations of children in twenty-three child-caring institutions were made. Of 2142 children tested, 143 were diagnosed as feeble-minded. The number tested, 2142, represents a fair sized sample of the 35,000 children in orphan asylums iin New York State. If less than 7% of orphan asylum children are feeble-minded, the situation is not so grave as has been pictured, but it implies the presence in the State of 2100 feeble-minded probably dependent children. It should not be assumed that the remaining 93 or 94% of orphan asylum children are fully normal, as many of them are retarded, dull, subnormal, or border-line cases. The heredity background of such a group of children is one which naturally presents itself for investigation by the sociologist. The families of all such children, including American types of parents, have been found more difficult to analyze than naturally would be supposed. The reason for this is that they have not realized that their child is defective or that he is a special problem. If the child does not do well in school work it is not recognized as an important matter by parents who are themselves more or less illiterate.

The study illustrates the need for special groups in schools and children's institutions where scholastic training will be given in proportion to mental capability. Those who are markedly defective should be removed from the school and committed to special institutions. Such children as are allowed to leave school should have a suitable occupation found for them, should work outside only when on probation, which should be extended for an indefinite period. Those failing in this plan would get along in a suitable colony where they could to some extent enjoy the fruits of their labors without being allowed the right of reproduction.

Chester L. Carlisle, M. D.

Superintendent, Division of
Mental Defect and Delinquency.
New York State Board of Charities.

The Journal of Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

NOTICE TO SUBSCRIBERS. Notify the publishers, Warwick & York, Inc., of any change of address; as copies lost in transit because of subscriber's removal to another address will not be duplicated.

THE JOURNAL is issuing a series of Monograph Supplements for the publication of more extensive researches in educational psychology. For details prospective contributors may apply to H. E. Buchholz, 10 East Center St., Baltimore, Md.

The price of the JOURNAL is $3.00 per year of ten monthly issues, payable in advance. Single issues are 50 cents each. Back volumes are $5.00 in original wrappers and $6.25 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME IX NOVEMBER, 1918 NUMBER 9

EDITORIAL

PEACE PROBLEMS

The war is over. Victory is won. The kind of peace for which we have fought is assured. Now we are confronted with problems of re-organization and rehabilitation, upon the successful solution of which depends the future prosperity of the world. Old ideas and conceptions are in the crucible. The tremendous political and social upheavals of central and eastern Europe are having their reverberation in the mutterings of social unrest in western Europe and America. Having made the world safe for democracy we are now face to face with the task of developing a democracy that is safe for the world. In a broad but very real sense this is a task of education.

Our educational leaders are by no means oblivious to the present need. From the press, the pulpit and the platform comes the recognition that the future welfare of the country rests upon education. Within a fortnight two of our most conspicuous educators, Ex-president Eliot, of Harvard, and President Butler, of Columbia, have discussed the future needs of education at some length. Dr. Eliot reiterates and amplifies his well-known view that in our edu-

cation we should pay more attention to the creative and interpretive activities of hand and voice. The eye should be trained to see, the ear to hear, and the hand to execute. This should be brought about through more intensive studies of the world in which we live, a greater emphasis upon the natural and social sciences, the interpretation of history in the light of present needs, and a more adequate utilization of the fine and industrial arts. Particularly do we need to divest our educational procedure of much of the formalism with which it has become encrusted, and to vitalize it by making it both more utilitarian and more idealistic. The paramount task of present education is to formulate and render articulate the inchoate demands of the new social outlook, and to devise methods of procedure which shall attain most effectively and economically the desired ends.

According to Dr. Butler the war has burnt up much of the dross and crudity of our earlier educational mintage, has torn to shreds the false psychology and economics of the "superman," and has reaffirmed the supremacy of the moral ideal. Further development of the moral ideal calls for training in ethics, economics and politics, as the fundamental verities in education. "There will be much more attention paid to the determination of individual differences of taste and capacity, and to making provision for them. The object of this determination is to prevent waste of effort, the loss of opportunity, and the blunting of talent by trying to sharpen it upon the wrong whetstone." Old methods of teaching come in for much criticism, particularly the teaching of physics, chemistry, biology and foreign languages. The elementary school is roundly secored for neglecting its proper business of thorough grounding in the fundamentals, and of running off after faddists and enthusiasts.

All of this is very interesting, but decidedly vague and sweeping. In regard to many aspects of our educational procedure, whether elementary, secondary, or college, we are in precisely the same position in which President Wilson recently confessed to Congress he found himself with reference to the railways: We do not know. Let us then frankly adopt the attitude of inquirers, both as to subject matter and as to method. Let us call for tests, as objective and scientific as may be, of what is being done and what is proposed. Let us demand, and let us train the public to demand, more experimentally verified fact and less dogmatic personal opinion. Let us grapple with the educational problems of peace as intelligently, as enthusiastically, as whole-heartedly as we did those of war. But let us above all things, in season and out of season, demand that we " be shown."

J. C. B.

NOTES AND NEWS

A preliminary meeting of the New York Society for the Experimental Study of Education was held at the Washington Irving High School, Friday evening, October 18. The general topic of the meeting was "What are the problems in education that are in most urgent need of experimental investigation, and how can this society proceed most successfully in attacking these problems?" The following papers were presented: "The Value of Experimental Studies for the Superintendent," Associate Superintendent Gustave Straubenmuller; "Opportunities for Experimental Study in Secondary Schools," Principal Arthur M. Wolfson, High School of Commerce; "Experimental Studies from the Principal's Point of View," District Superintendent Arthur C. Perry, Jr.; "The Teacher's Gain from Experimental Studies," Professor Thomas H. Briggs, Teachers College.

The council of the American Psychological Association first decided to abandon the annual meeting scheduled for December, 1918, on account of the exigencies of the war and the preoccupation of the members with war work. The signing of the armistice has led the council to reconsider its decision, and a brief meeting will be held at Baltimore, December 27 and 28, for the consideration of topics growing out of the war.

The temporary officers of the American Association of Clinical Psychologists at first deemed it advisable to cancel the annual meeting scheduled for December, 1918. After the cessation of hostilities, however, the plans were again changed, and a brief meeting will be held in Baltimore. The temporary officers of the Association are as follows: Chairman, J. E. Wallace Wallin; Secretary, Leta S. Hollingsworth; Committee on Constitution, Leta S. Hollingsworth, David Mitchell and Francis N. Maxfield; Committee on nomination of officers and new members, Rudolf Pintner, Helen Thompson Woolley and H. H. Goddard; Committee on program for the next meeting, the Chairman, secretary and the committee on constitution.

The summer training school of psychiatric social work conducted by the Boston Psychopathic Hospital and Smith College under the auspices of the National Committee for Mental Hygiene opened at Smith College July 7th with an enrollment of 68 young women from 21 states and as many colleges. The purpose of the school was to give in eight weeks the theoretical background necessary to prepare social workers to assist in the rehabilitation of soldiers suffering from "shell shock" and other nervous and mental disorders. The list of instructors included many well-known names, and a wide range of subjects was given which covered the field very thoroughly. Of the 58 students who completed the work satis-

factorily, those who had not previously had the required amount of practical work were assigned to various clinics and hospitals for further training.

The committee on educational psychology of the Society of College Teachers of Education, Professor Frank N. Freeman, chairman, has prepared and distributed a questionnaire designed to bring out just what type of work is given in the various institutions under the name of educational psychology.

The Bureau of Cooperative Research of Indiana University, under the direction of Dr. Walter S. Monroe, has issued an announcement of plans for investigations in 1918–1919, and a list of standardized tests for sale by the Bureau. Among the projects noted are the further standardization of Monroe's reasoning tests in arithmetic, a study of common errors in arithmetic, reading and spelling, and the collection of effective methods of correcting errors in these subjects.

The Bureau of Educational Experiments, 16 West 8th Street, New York City, announces the issue of Bulletin No. IX, *Psychological Tests, Revised and Classified.* This bulletin is said to be a much more extensive collection of material than their former Bulletin No. VI, and to include publications to October 1, 1918.

The General Education Board announces the publication of the Gary Survey in eight parts, as follows: I. *The Gary Schools: A General Account*, by Abraham Flexner and Frank P. Bachman; II. *Organization and Administration*, by George D. Strayer and Frank P. Bachman; III. *Costs*, by Frank P. Bachman and Ralph Bowman; IV. *Industrial Work*, by Charles R. Richards; V. *Household Arts*, by Eva W. White; VI. *Physical Training and Play*, by Lee F. Hanmer; VII. *Science Teaching*, by Otis W. Caldwell; and VIII. *Measurement of Classroom Products*, by Stuart A. Courtis. The parts vary in length from 34 to 350 pages and cost from ten to thirty cents. They will be published at brief intevals beginning December 1.

Dr. S. L. Pressey, of Indiana University, has devised a group scale for the measurement of intelligence. The tests consist of rote memory, logical selection, arithmetic problems, opposites, logical memory, word completion, moral classification, dissected sentences, practical information, and analogies. Tests of over 1000 school children indicate a good correlation with the Stanford-Binet scale and with the judgments of teachers.

The State Board of Health of California has created a bureau of child hygiene to study the child health of the state. Miss Amy D. Steinhard has been placed in charge of the bureau.

Dr. Lotus D. Coffman, dean of the school of education of the University of Minnesota, has been appointed head of the work by the War Department on the reconstruction of crippled and injured soldiers, succeeding Dean James E. Russell, of Teachers College, Columbia University.

Professor E. A. Bennett, dean of Highland College, Des Moines, Iowa, has accepted the chair of education at Boston University.

Dr. Joseph Peterson, assistant professor of psychology in the University of Minnesota, has resigned to accept a professorship in psychology in George Peabody College for Teachers, Nashville, Tenn.

Miss Frances Lowell has been appointed instructor in psychology at the University of Minnesota.

Dr. Walter B. Swift, of Boston, has been appointed consultation expert for speech defects to the division of medical inspection of the public schools of Cleveland, Ohio.

Mr. George A. Mirick, formerly deputy commissioner of education in New Jersey, has been appointed assistant professor of education at Harvard University.

Professor Lightner Witmer, head of the department of psychology at the University of Pennsylvania, has returned from his work with the Red Cross in Italy, and is on active duty this year.

Captain Madison Bentley is now president of the Aviation Examining Board, examining aviation recruits from New England colleges.

Dr. Roswell P. Angier, professor of psychology at Yale University, is a captain in the Sanitary Corps, National Army, at the Hazelhurst Field Medical Research Laboratory, Mineola, L. I.

Dr. Ernest N. Henderson, professor of philosophy and psychology at Adelphi College, has been commissioned captain in the Sanitary Corps.

A new volume by Professor Thorstein Veblen, of Cornell University, entitled *The Higher Learning in America: A Memorandum on the Conduct of the Universities by Business Men*, is announced for early publication by B. W. Huebsch.

Among the forthcoming books announced by Longmans, Green and Company are *The New State*, by M. P. Follett, which presents group organization as the solution of problems in popular government; *The Citizen and the Republic*, a textbook in government, by James Albert Woodburn and Thomas Francis Moran; and *The Human Machine and Industrial Efficiency*, by Frederick S. Lee.

PUBLICATIONS RECEIVED

JAMES ROWLAND ANGELL. *An Introduction to Psychology.* New York: Henry Holt and Company, 1918. Pp. vii, 281.

This is not a mere condensation of the author's older and larger work, but a fresh survey of the entire field of psychology. "The *organizing* character of every level of our psychological activities is more explicitly brought out than in the older book, with advantage, I trust, to the depth and vividness of impression left on the reader's mind regarding the real nature of our adjustments to environment." Without going to the extreme of rejecting introspection entirely, the author manifests a strong leaning toward the behavioristic point of view. Particularly the doctrine of imageless thought and of the influence of subconscious neural activities upon the stream of consciousness are regarded with increasing favor. It is noteworthy that the ancient, atomistic conception of instinct is still dominant, and that Thorndike's vigorous onslaught on this doctrine has had little effect on the discussion. As a simple introductory outline the book will undoubtedly prove highly serviceable.

HERBERT BATES. *English Literature.* New York: Longmans, Green and Company, 1918. Pp. xiv, 605 $1.50.

What a striking contrast is this book with the dull, sombre histories of English literature of our early memories! Here we have an abundance of illustrations from old prints or the reproductions of great masterpieces, many choice selections illustrating the excellencies which the author extols, numerous diagrams, a fascinating literary map of England, and an abundance of helpful guides to further reading. The text itself tells a delightful story and should make the study of English literature as absorbing as it formerly was stupefying.

The Boston Way. Plans for the Development of the Individual Child. Compiled by the Special Class Teachers of Boston. 1917. Pp. 127, $1.00.

This little book is packed with valuable suggestions for teachers of backward or defective children. The forty-two school subjects included range from arithmetic, reading and spelling to cobbling, crocheting and brush-making. True, the psychology is still in the faculty stage, and sense training is taken for granted, but the incidental theory does not at all nullify the value of the ingenious and diversified practical recommendations.

JEAN BROADHURST. *Home and Community Hygiene.* Philadelphia: J. B. Lippincott Company. 1918. Pp. xiii, 428. $2.00.

This is one of the excellent Lippincott's Home Manuals series under the editorship of Benjamin R. Andrews, of Teachers College. The sub-title is "A Text-book of Public and Personal Health." There is a distinct need for a book that brings together in popular, untechnical language the latest results in the science of healthful living. The present volume performs this service admirably and deserves a very wide circulation.

HENRY REED BURCH AND S. HOWARD PATTERSON. *American Social Problems, an Introduction to the Study of Society.* New York: The Macmillan Company, 1918. Pp. ix, 381. $1.20.

It is remarkable how the social sciences, the sciences of man, have been neglected in our secondary curricula. We talk much of our education as a preparation for

living, yet we avoid any systematic effort to make the bases of social life comprehensible. The present book aims to present in very elementary fashion some of the essential features of our social organization. From the point of view of the evolutionary conception the development of civilized from primitive man is sketched, the influence of climatic and physiographic features is indicated, the development of the family and the state is outlined, and such problems in American social life are discussed as immigration, the negro race, industrial adjustment, poverty, charity and crime, the evils of mental defect, drink, and divorce, and the function of the school and other social agencies in moral progress. The book should stimulate high school students to reflection and wider reading on social questions.

ARTHUR D. DEAN. *Our Schools in War Time and After.* Boston: Ginn and Company, 1918. Pp. v, 355. $1.25.

The problem of how we can utilize the energies and activities mobilized for the war to the betterment of the training of our children for lives of peace is occupying the attention of all thoughtful students of education. The first requisite is a survey of the work that our schools have done during war time, and this is admirably presented in the volume before us. There is an account of the response of our schools to war needs, the services rendered by industrial and trade schools, colleges and technical institutes, war activities in manual and household arts, organizing boy power, red cross, re-education of the disabled, farm cadets, and a program for the future. The discussion points the way to many modifications of school procedure that will be useful in peace as well as in war.

MARY DUCLAUX. *A Short History of France from Caesar's Invasion to the Battle of Waterloo.* G. P. Putnam's Sons, 1918. Pp. ix, 345. $2.50.

The author states that she has written this book neither for schoolboys nor historians, but for the great mass of cultivated but ignorant men and women who, particularly at the present time, desire to know more of the past of that heroic sister republic to whose aid we have given of our best. She has given us a vivid, stirring tale in which the great figures of French history stand out with lifelike distinctness. It is not a social, political or industrial history so much as a history of striking personalities. It is this character that lends the book its enthralling interest.

ABRAHAM FLEXNER AND FRANK P. BACHMAN. *The Gary Schools, A General Account.* New York: General Education Board, 1918. Pp. vi, 265. Twenty-five cents.

This is the first of the reports on the survey of the Gary schools by the General Education Board. It gives the general plan of the survey, summarizes the important findings that will be presented in greater detail in subsequent reports, and draws conclusions as to the significance of the survey as a whole. Salient features of the latter are the credit due to Gary for adopting a modern and progressive school policy, the plan of organization for presenting an enriched curriculum at a minimum expense, democratic practice in school discipline, lack of watchful administrative control, defective supervision, and some poor teaching.

GRACE FULMER. *The Use of the Kindergarten Gifts.* Boston: Houghton Mifflin Company, 1918. Pp. vii, 232.

This discussion of fundamental principles of kindergarten education is founded firmly on the modern psychology of adjustment. It is an attempt to interpret the traditional materials of the kindergarten in terms of modern ideas and to free them from the mystical obscuration of the orthodox Froebellian exposition. The attitude toward Froebel is sympathetic not critical, and the book has grown out of long experience in the kindergarten and in training kindergarten teachers.

WILLIAM ERNEST HOCKING. *Human Nature and its Remaking.* New Haven: Yale University Press, 1918. Pp. xxvi, 434. $3.00.

This is one of the most profound and significant books of this generation, and yet withal so interestingly and attractively written that it reads like a fairy tale. It is an attempt to establish a system of moral and social values upon a consistent biological basis. The units of this biological structure are the instincts. The author acknowledges that the conception of instinct is purely hypothetical, that the examination of any specific bit of human behavior reveals a complex of habits wrought out under the pressure of experience from the primitive materials of inherited structure and function. The argument shows that he is perfectly familiar with Thorndike's searching analysis of these hypothetical units. Nevertheless, he finds it convenient for his purpose to assume the existence of these composite structures of heredity and experience, and then proceeds to show how there grow out of these native tendencies such complex attitudes as the will, conscience, sin, social and political obligation, and religious beliefs and aspirations. The book might well be called a natural history of social and moral control. It is of tremendous significance for the theory of education.

E. M. JAMISON, AND OTHERS. *Italy, Medieval and Modern.* New York: Oxford University Press, 1917. Pp. viii, 564. $2.90.

The part that Italy has played in the recent world-war arouses an interest in her development which will make the present volume very acceptable. "Our historical literature is full of admirable monographs on particular epochs of Italian history . . . . Yet it is strangely difficult to find any general sketch of Italian history, from the barbarian invasions to the present day, which can be recommended as an introduction to more detailed studies. It was to supply this need that the present volume was planned and written."

CHARLES HUBBARD JUDD. *The Evolution of a Democratic School System.* Boston: Houghton Mifflin Company, 1918. Pp. ix, 119.

"The feeling has been steadily gaining strength that our generation must shake off the institutional traditions of a past age and organize a sound scheme of democratic education. The present study is an effort to bring out explicitly some of the justifications for the reorganizations which are now under way. The book aims to bring to clearer consciousness the unique characteristics of our continuous educational system. It aims to point the way by which much of the present waste of pupil's time and energy can be corrected. It is a plea for a tolerant attitude toward the crudities of the junior high school. It is a plea for more co-operation in developing this institution."

CHARLES HUBBARD JUDD, AND OTHERS. *Reading: Its Nature and Development.* Chicago: The University of Chicago Press, 1918. Pp. xiv, 192. $1.00.

The present monograph gives a summary of the extensive investigations in reading which have been carried on at the University of Chicago for the past two or three years with the financial support of the General Education Board. The chapters deal with changes in the teaching of reading, a study of reading books, eye-movements of adult readers, analysis of children's reading, special training in reading, progress through the grades, individual differences, and reading for meaning. The author urges greater emphasis on reading for meaning in the upper elementary grades.

EDGAR JAMES SWIFT. *Psychology and the Day's Work.* New York: Charles Scribner's Sons, 1918. Pp. ix, 388. $2.00.

This book might aptly be called the browsings of a psychologist. The author takes his texts from the newspapers, from literature, from business, from the war, from the history of science, from the frailties and foibles, the grandeur and heroism of man wherever he finds them, and out of these develops the most entertaining reflections. The most significant chapters for the teacher are those on the psychology of learning, curiosities of memory, memory and its improvement, and the psychology of testimony and rumor. While there is no systematic exposition of any particular psychological creed, the author's psychology is always sound and rests upon well established experimental data. Indeed the author's avowed purpose is to present the essence of experimental psychology in such untechnical and attractive form that the layman can and will become acquainted with it and profit by it.

LEWIS M. TERMAN. *A Trial of Mental and Pedagogical Tests in a Civil Service Examination for Policemen and Firemen.* Reprinted from the Journal of Applied Psychology, 1: 1917, 17–29.

The tests used were the Trabue Completion Tests, the Thorndike Reading test, the Ayres scale for handwriting, a spelling test based on Starch, and the Courtis Arithmetic, Series A. The Stanford revision of the Binet tests was used. The intelligence quotients varied from 60 to 114, median 84. The highest correlation was between mental age and arithmetical reaoning (.81).

CHARLES W. WADDLE. *An Introduction to Child Psychology.* Boston: Houghton Mifflin Company, 1918. Pp. xvii, 317. $1 50

This book is a valuable addition to the growing literature on child psychology. The first five chapters give the historical and biological background of child psychology, chapters six to nine deal with play, linguistic development, children's drawings, and the moral development of children. chapter ten discusses juvenile delinquency, and the last two chapters present the general facts and principles of mental development and the unfolding of specific capacities. Each chapter is followed by a summary, a list of questions, and a fairly extended bibliography. It is strange that the author in his chapter on "Non-learned Human Behavior" quotes extensively from McDougall, Hall and others, but almost entirely ignores the critical work of Thorndike.

HOWARD C. WARREN. *The Mechanics of Intelligence.* Reprinted from the Philosophical Review, 26, 1917. Pp. 602–621.

The central problem of this paper is how does intelligence act on voluntary muscles. An interesting illustration in the course of the discussion is an analysis of the controls involved in playing a game of chess. The author's conclusion is that intelligence has no causal significance whatever, but is merely the accompaniment of neural activities which are as mechanical in their nature as other physical and chemical changes.

ARLAND D. WEEKS. *The Psychology of Citizenship.* Chicago: A. C. McClurg and Company, 1917. Pp. 152. $0.50.

Since citizenship demands intelligence any study of the activities of intelligent citizenship may have a certain right to be called psychology. Of psychology in the modern critical, analytical sense this book contains little. Yet it is a good book, stimulating, thought-provoking, liberalizing. The chapter on the control of suggestion, for example, gives a splendid account of the factors in contemporary life which exert effective control over our conduct.

DAVID EMRICH WEGLEIN. *The Correlation of Abilities of High School Pupils.* The Johns Hopkins University Studies in Education, No. 1. Baltimore: The Johns Hopkins Press, 1917. Pp. 100. $1.25.

This is a careful study of the relationships found in the high school marks of a group of 255 girls in the subjects of the course of study taken separately by years. There is first a good summary of previous studies in correlation, than an examination of each year's records in detail, and finally a report on three mental tests given to certain members of the group.

Whittier State School. Biennial Report. Whittier, California, 1917. Pp. 245.

Sixty-seven pages of this volume are devoted to the report of the Department of Research, which is under the directorship of Dr. J. Harold Williams. Among the topics discussed are the nature and scope of research in such an institution, a popular account of the measurement of intelligence,heredity and delinquency (with many valuable case studies), environmental factors in delinquency (with the application of the Whittier Scale for Grading Home Conditions), and problems in the study of delinquency. The report of the education department gives the results of several standard tests. The report as a whole sets a high standard of scientific and typographical excellence.

J. HAROLD WILLIAMS. *Delinquency and Density of Population.* Reprinted from the Journal of Delinquency, 2: 1917, 74–91.

The evidence shows that the popular belief that crime is more characteristic of cities than of the country is an error. Delinquency is most prevalent in small towns. There is no particular class of offenses associated with cities, towns or rural communities. The intelligence of delinquent boys from cities is higher than that of boys from the country.

LEON L. WINSLOW AND AUGUST P. GOMPF. *The Industrial Arts in Elementary Education.* Bowling Green, Ohio: Published by the Authors, 1917. Pp. 110.

Disciplinary manual training has had its day, and now gives place to the industrial arts. This suggestive little outline for the utilization of industrial arts material will be a boon to progressive elementary teachers. The topics include stone, clay and glass products; metal products; wood manufactures; chemical products; paper products; books and printed products; textiles and clothing; and foods. There is a list of products for each grade, a classification of industrial products, and a list of helpful reference books.

HERBERT WOODROW. *The Faculty of Attention.* Reprinted from the Journal of Experimental Psychology, 1: 1916, 285–518.

The author has chosen his title deliberately. He contends that as a result of extended experimentation with touch, sound and light re-actions he cannot escape the conviction that "every individual has a certain power or faculty of attention, in the sense that the degree of his attention is determined in part by general conditions which remain effective in spite of variation in the specific type of the mental activity in question." But this seems to be nothing more marked than the "learning type," wherein we find some people fast learners and others slow, no matter what the material they are learning. Yet one does not ordinarily hear of "the faculty of learning."

C. S. YOAKUM AND MARY C. HILL. *Persistent Complexes Derived through Free Associations: Miss Z's Case.* Reprinted from the Journal of Abnormal Psychology, October, 1916. Pp. 45.

This is a case of remarkable facility in remembering the Binet letter squares, in which it developed that the subject made use of highly fantastic free association stories to hold the letters in their order. Examination of these stories led to the discovery of peculiar suppressed complexes that had a strong influence on the subject's general reactions.

C. S. YOAKUM AND MARY C. HILL. *Genetic Antecedents of Free Association Materials: Miss Z's Case.* Reprinted from the Journal of Abnormal Psyhology, February, 1917. Pp. 15.

An account of the previous history of the subject described in the earlier paper, and an explanation of the origin of the tendency to invent "stories" to assist in meeting situations.

VOLUME IX DECEMBER, 1918 NUMBER 10

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

MEASUREMENTS OF SOME ABILITIES IN SCIENTIFIC THINKING

JOHN P. HERRING
Seattle, Washington

NEED.—Civilized society needs more scientific ability.

There is a most vexing scientific problem which arises daily in all our lives. Some circumstances there are amid which it is wise to rely upon competent authority; others, upon personal investigation; and yet others, upon both. How may our population distinguish among these? Amid what circumstances trust conclusions? Even slightly keener insight into this problem is valuable.

A distinction, vital, yet forming no defined part of these measurements has here to be made,—a distinction, in the sense to be defined, between primary and secondary data. Primary data are selected by the individual from immediate experience; secondary data, from representations of the immediate experience of other individuals. When I see James board the street-car, I am selecting a primary datum. When John tells me that James boarded the street-car, I am selecting a secondary datum. The primary are always the basis. To them is the final reference. They are more educative than the secondary. Their use in our schools is still too rare. Nothing so characterises the modern period of history as its insistence upon building its thought world when it can out of primary data, and its knowledge that secondary data are subjected consciously and unconsciously to certain streams of influence, such as prejudice and carelessness, not to mention greed, from which the sole escape is vigilant comparison with primary data. We in the school are still too exclusively mediaeval schoolmen, in the sense of depending too much upon secondary data for our conclusions, and of being too content to encourage our students to conclude upon the bases of secondary data alone. Too often we accept statements from books instead of those from life that could be found easily and used

convincingly. To let print pass as evidence when experience is both available and educative is to revert to attitudes which it was a function of the Renaissance to replace. For the middle ages Aristotle was a final authority; for this modern age no authority is final. In this age we must train children to rely much upon the testimony of their own senses, taken with their interpretation of the material furnished them by their senses. It is not scientific to rely habitually upon the conclusions of others without seeking primary data, for which the search is often feasible. This emphasis has been placed by some writers and teachers, but it is general neither in thought nor in practice.

When and how to investigate, when and how to recognize authority, when and how to combine investigation with authority,—these are problems of scientific method, of education, and of citizenship. It is sound educational policy to give them central importance in school. How many dismal errors of judgment at critical moments in public and private life could thus be averted,—errors in the choice of political candidates, of employment, of studies, of vocation and avocation, of national and individual friends and enemies, of dress, of residence, of attitudes, of schools, of courses of study, of food, clothing, shelter, education, and amusement,—of all,—everything,—that ever makes or mars our lives?

AIM.—The ultimate aim of this investigation is to aid in focusing attention sharply upon scientific method as the proper method of children's study, both in the eleven abilities tested, and in all other abilities which are at once important and feasible.

The immediate aim of this investigation is to increase the consciousness and effectiveness of the use of eleven scientific abilities.

Scientific ability is used daily both by adults and by children. It is used more or less consciously, accurately, effectively, and appropriately. When pupils make good use of method, their world of data and their education are ordered and relatively useful. The scientific method is the best method. In its highest reaches it is complexity itself. In its simplest forms its elements are employed unconsciously by children in the pre-school age. The published mathematical calculations of the position of the planet Neptune, before it had ever been seen, filled several volumes. A girl of three said: "Big engine says, 'Ding dong;' 'Pussy's in the Well' says, 'Ding dong;' many things say, 'Ding dong.' " She collected two facts and made an inductive conclusion. A child of five, who said,

"A pond is a little round place to sail boats in—with water in," defined imperfectly; nevertheless, to a certain extent, she employed ability number three. Her mental work falls within the concept of scientific method. It is such crudely elementary uses of method by the masses of children upon which we must erect the superstructure of more advanced and finished work. We must teach no less of the content of geography, but we must lay more stress upon teaching children to study, to handle problems, to work methodically, to plan independently their own ways of solving problems, to lay out their own work, to observe precisely, to remember exactly and fully, to think well, to expound clearly,—in short, to use the scientific method.

The ultimate aim of this investigation is wider than the present scope of the test.

DEFINITION.—The competent have never formulated even a tentative concensus of judgment in answer to the question: What is scientific method? In the absence of such an agreement concerning the living and developing complex of processes known as scientific method the following is a more vague and arbitrary statement, of which the first part expresses a general meaning, and the second a meaning narrowed by analysis and selecton to the particular purposes of this investigation.

The aim of science is to explain the world. The aim of *a* science is to explain a limited portion of the world. The scientific method is that complex of processes used in the attempts to explain. It is a complex of processes which are never mutually exclusive. It is a complex into which its elements enter in ever varying frequency, order, and degree. While it can be analyzed usefully, it can be analyzed neither finally,—for scientists alter its use from age to age; nor universally,—for scientists alter its use from field to field.

The history of the theory of scientific method is suggested by such names as Francis Bacon, John Stuart Mill, W. Stanley Jevons, and Karl Pearson.

ANALYSIS.—From among the processes of which the total complex is scientific method, eleven, which have importance with reference to scientific work, and feasibility with reference to inclusion in this scientific method test, have been selected. While all of them together are very far from constituting the whole of scientific method, severally they fall quite within that concept. They are called more or less into play whenever a problem is to be solved carefully,—that

is, whenever the scientific method is to be used. Each ability is a process used in science. In the list following, each is designated by a catch-word. Each catch-word appears at the top of each of three exercises. Each exercise is intended as a partial measure of one ability. (The catch-words do not appear in the test as given to the students.) Following this list is an equal number of paragraphs, each devoted to the definition of one ability.

1. *Value*
2. *Feasibility*
3. *Definition*
4. *Clarity*
5. *Statistics*
6. *Relevancy*
7. *Recording*
8. *Comparison*
9. *Classification*
10. *Arrangement*
11. *Sufficiency*

1. *Value.* The ability of pupils to judge as to the relative values of different problems. Whenever a problem is suggested for solution, this question should arise first: Is the problem worth solving? Will its solution be of enough value to pay for the work it will take? Will the conclusions of the investigation be of worth?

2. *Feasibility.* The ability of pupils to judge whether problems can or cannot be solved. When one has decided that it would be valuable to solve a certain problem, the very next question is: Can it be solved? Can suitable methods be devised? Can enough relevant data be found?

3. *Definition.* The ability of pupils to distinguish good definitions from bad. When one has decided to try to solve some problem, the need of defining some of its terms often arises. Pupils need to know the differences between good and bad definitions. A definition of a thing usually states a class in which the thing belongs, and a mark or marks by which the thing may be distinguished from differing things in the class.

4. *Clarity.* The ability to distinguish sentences which contain no ambiguity from those which contain ambiguity; to locate ambiguity; to recognize clarity of expression.

5. *Statistics.* The ability to know when statistics are needed in the proof of a statement or in the solution of a problem.

6. *Relevancy.* The ability to select facts that relate to the problem and to reject those that do not; to stick to the point. When one has chosen his problem and made clear its meaning, one must know how to select data that will aid in its solution, and to reject those that will not.

7. Recording. The ability to select the better from among the poorer methods of recording the facts relevant to a problem.

8. Comparison. The ability to distinguish good from bad comparisons. Comparison involves the likenesses and the differences of two or more things.

9. Classification. The ability to distinguish between a class containing something foreign to it and a class not containing something foreign to it; to recognize the presence of something foreign in a class.

10. Arrangement. The ability to arrange the members of a class in useful sequence; to distinguish between good and bad arrangements. When one has placed the relevant objects or data each in its own class in accordance with some purpose, the members of each class should usually be set one after the other in some sequence. The sequence may be dominated by any one of a great variety of ideas, according to the purpose, such as number, size, color, or degrees of importance, of difficulty, or of obviousness.

11. Sufficiency. The ability to judge whether the data are sufficient for the purpose. Having the relevant data thus classified and arranged, no conclusion may be drawn until the question of sufficiency has been answered: Are the data of enough combined weight to prove the point? Are they adequate to solve the problem? One method of proof is never considered sufficient for a proposition of much difficulty or importance, but must, when possible, be supplemented, that is, substantiated, checked, verified, by one or more other methods. Only a verified proof is sufficient. Verification is somewhat the same idea as sufficiency.

The test as given to students is printed on an eight page folder in eight point type. On the first page are blanks for name, date, address, grade, class, teacher, scores, and a sample exercise which is explained and solved. There are also the following:

Directions to Students

These exercises are given to see how well you can think. On the following pages you will find other exercises like the sample above. Do the first, then the second, and then the others in order, until you have done all you can. You will have all the time you need. Work all the time. Look over each exercise, after you have done it, to be sure you have done it right. If you make any mistake, correct it. Ask no questions. Let no one help you.

After you have done all the exercises you can, take them to the teacher at once. Take your pencil with you. As soon as the teacher has written down how long it took you to do the test, write the answers to the two questions below:

Did you get any help on any exercise on the following pages? Answer "Yes," or "No."................

If you did, what help did you get?.................................... ..

The Test

1. Classification

Put a cross on the line before the word that is in the wrong group of words; but if there is no such word, put a cross here............

Animals	*Buildings*	*Foods*	*Fruits*
........cat	........house	........meat	........apple
........cow	........barn	........bread	........dog
........horse	........store	........butter	........cherry

2. Clarity

Put a cross on the line before the sentence that is not clear; but if there is no such sentence, put a cross here........

........Some lakes are small, some lakes are large, and some are very large.
........Some ponds are small, and some are large.
........Men are not all the same height, for some are taller than others.
........Some men are short, and some men are very tall.
........Some women are short, and some women are very tall.
........Women are not always of the same height as men.
........Many ponds are small.
........Many lakes are large.
........John and James went fishing, and he was drowned.
........Some mountains are larger than others.
........Some lakes are larger than others.
........Some men are taller than others.

3. Relevancy

Put a cross on the line before the fact that helps show that the air is colder high up than it is low down; but if there is no such fact, put a cross here.........

........Tables have four legs, braces, and tops, and are very useful for many purposes.
........Chairs have four legs.
........Some houses are colder than others.
........Cradles sometimes rock.
........Some men are colder-blooded than others.
........The animals give us much food and clothing.
........Horses are larger than dogs.
........Mountain peaks are cold.
........Animals usually have ears.
........Animals usually have eyes.
........Some animals are colder-blooded than we.
........Some chairs rock.

4. Definition

Put a cross on the line before the best way of telling what a pond is; but if there is no such way, put a cross here........

........A thing with water all the way around it.
........Something useful to boys and girls who like to swim.
........Something surrounded by people who are talking.
....X...A small body of water surrounded by land.
........A house surrounded by a large forest in this country.
........A street.
........A number.
........A book.
........Lands.
........Water and land.
........A body of water, large or small, and almost surrounded by land.
........A body of land.

5. Feasibility

Put a cross on the line before the statement that would be too hard to prove true; but if there is no such statement, put a cross here.........

........Some children are taller than others.
........Some women are older than others.
........Some old men are smaller than others.
........Some children are larger than others.
........Some children have better homes than others.
........Some children know more than others.
........Some children know less than others.
.....X..The brightest child lives in Boston.
........Children go to school, at least in this part of the world.
........Grown people do not.
........Some men know more than others, and the same is true of women.
........Some men know more than some women.

6. Arrangement

Put a cross on the line before the group of things that come in the best order; but if there is no such group, put a cross here........

.....x..baby, boy, young man, old man
........baby, young man, old man, boy
........old man, boy, baby, young man
........baby, boy, old man, young man
........old man, young man, baby, boy
........boy, old man, young man, baby
........boy, baby, young man, old man
........young man, old man, boy, baby
........young man, baby, boy, old man
........boy, old man, young man, baby
........baby, boy, old man, young man
........old man, young man, baby, boy

7. Clarity

Put a line under the word that makes the following sentence not clear; but if there is no such word, put a cross here.........

Boston and New York are both large cities, but it is smaller than the other.

8. Recording

Put a cross on the line before the best way of keeping facts in geography until you are ready to use them; but if there is no such way, put a cross here........

........Write them down, put the paper in your book, and lose the book.
........Talk plainly to someone who knows how to keep the facts he will want to use.
........Write your name.
........Write a friend's name, then your mother's, and then your father's.
........Speak.
........Speak loud.
........Go somewhere.
....X..Write the facts down, using just a few words for each.
........Write.
........Write your name fast; then write it again and again without stopping.
........Write your name slowly.
........Write your name well; then try to write it better and better.

9. Sufficiency

Put a cross on the line before the way by which you could surely find out how much rain falls in New York City in a year; but if there is no such way, put a cross here........

........Ask any one you meet.
........Think about it hard.
.....X.Look in the right book.
........Ask some friend who lives far away.
........Ask two friends who live near.
........Guess at it.
........Guess at it and have two or three friends guess at it.
........Sit still.
........Do not whisper; always do what your teacher tells you to do.
........Be good.
........Sit still and think hard.
........Write about it to your friends.

10. Arrangement

Put a cross on the line before the best order of things; but if there is no such order, put a cross here........

........cloud rain river brook ocean
........rain cloud brook river ocean
........ocean rain cloud brook river
........river ocean rain cloud brook
........cloud brook river ocean rain
........brook river ocean rain cloud
........brook rain cloud river ocean
........rain cloud river ocean brook
........cloud river ocean brook rain
........cloud rain brook river ocean
........river ocean brook rain cloud
........ocean brook rain cloud river

11. Value

Put a cross on the line before the question that is most important to the greatest number of persons; but if there is no such question, put a cross here........

........Is your father ill?
........Does Mr. Jones own his house?
........Is your pencil sharp?
........Can you run a mile without stopping?
........Will the earth continue to have enough light?
........Does a pound of feathers weigh as much as a pound of lead?
........Does candy ever cost eighty-five cents a pound?
........Does candy ever cost thirty-five cents a pound?
........Are some fire-places built of brick?
........Are some of the very best stoves made of iron?
........Have tables four legs each?
........Do some chairs rock?

12. Definition

Put a cross on the line before the best way of telling what a pond is; but if there is no such way, put a cross here........

........A body of water wholly surrounded by land.
........A body of water surrounded by land.
........Water which is of use to boys and girls.
........Water surrounded by land.
........A small body surrounded by land.
........Still water.
........A small body of water surrounded by land.
........A body of land surrounded by water.
........A large body of water surrounded by land.
........A small body of land surrounded by water.
........A body of land fitted for man to live on.
........Land and water.

13. Feasibility

Put a cross on the line before the statement that can be proved true; but if there is no such statement, put a cross here.........

........The telephone is just as useful as the telegraph.
........The heat of summer is harder to bear than the cold of winter.
........It is better for all persons to live in valleys than upon hills.
........Next summer will be hotter than last summer.
........There is always too little rain.
........Dogs never do harm.
........Iron-making is the most useful industry.
....X..Canada is farther north than the United States.
........Sometimes two persons are just alike.
........Sometimes two blades of grass are just alike.
........A pound of lead weighs more than a pound of feathers.
........Sometimes two Americans are just alike.

14. Relevancy

Put a cross on the line near the statement that helps show that air is colder high up than it is low down; but if there is no such statement, put a cross here.........

........Too little rain hurts plants.
........Men can sail in the Arctic Ocean only a small part of the year.
....X..There is snow in summer on some mountains.
........Plants grow fast in warm, wet places.
........Dogs have saved many persons lost in the snow.
........The North Pole was discovered.
........It is cooler in the shade than in the sun.
........Snow lasts longer in the shade than in the sun.
........There are railroads on the sides of some high mountains.
........Men should be strong to climb some high mountains.
........Too great cold hurts plants.
........Frost hurts some plants.

15. Comparison

Put a cross on the line before the sentence that shows a difference between two things; but if there is no such sentence, put a cross here........

........The Mississippi River is long, winding, and beautiful.
........Some rivers are short, rapid, and almost useless.
........The Mississippi River carries boats of many kinds.
........The boats carry cotton, wheat, and other things, besides passengers.
........China makes 15 million pounds of silk a year, Japan 7.
........Many boats carry passengers up and down the Mississippi River.
........Many of the boats have water-wheels to make them move.
........The water-wheels are placed at the ends of the boats.
........The water-wheels turn because they are driven by the engines.
........The turning makes the boats go.
........The boats have engines.
........The engines make the wheels turn.

16. Statistics

Put a cross on the line before the statement that could be more surely proved by using numbers than in any other way; but if there is no such statement, put a cross here........

........Some boys use boats well; some build boats well, too.
........Both girls and boys go to school in this country.
........Men and women in this country sometimes work to earn a living.
........Schools are made for teaching; they are sometimes used for public meetings.
........Children go to school to study.
........Children go to school to learn to study; they do not always succeed.
........There are teachers who do not teach children to study.
........There are beautiful children who go to school.
........This country contains some wildcats.
....X..A pound of sugar costs more than a pound of flour.
........Some houses are built of brick.
........A most important thing about school is learning to study.

17. Classification

Put a cross on the line before the best grouping of things; but if there is no such grouping, put a cross here........

........house lion wildcat cat   tiger school-house store
........house store cat lion   tiger store wildcat
........lion cat store house   tiger store wildcat
........tiger store wildcat   lion cat store house
........tiger store wildcat   house store cat lion
........tiger lion wildcat cat   house school-house tiger
........house wildcat store   tiger store wildcat
........house store wildcat   tiger wildcat store
........lion house cat store   tiger wildcat store
........house school-house store   tiger lion wildcat cat
........tiger wildcat store   lion store house cat
........house store cat lion   tiger wildcat store

18. Value

Put a cross on the line before the question that is most important to the greatest number of people; but if there is no such question, put a cross here........

....X..Will the sun continue to rise every morning?
........Are evergreens found in many parts of the earth?
........Is coal found in Alaska?
........Do rivers always flow from higher to lower ground?
........Do some mountain ranges run from north to south?
........Do some mountain ranges run from east to west.
........Was tobacco discovered in this country?
........Has the telephone been used for many years?
........Is coal mined in this country?
........Does water freeze in very cold weather?
........Does the moon go around the earth in about twenty-eight days?
........Is oil found in the earth with coal in this country?

19. Value

Put a cross on the line before the question that is most important to the greatest number of people; but if there is no such question, put a cross here........

........Does man raise enough wheat?
........Do the forests yield enough wood?
........Do the strongest nations have the best schools?
........Do our cows and goats yield enough milk?
........How much coal is there left in the earth?
........Are the fur-bearing animals in danger of being killed in too great numbers?
........Are inland countries well connected with the sea by means of rivers?
........Are the continents well connected with each other by means of the oceans?
........Will the earth continue to have a good supply of heat?
........How much is the telephone used in the country?
........In what languages are the best books likely to be printed?
........How many more automobiles have been sold during the last year than during the year before that?

20. Statistics

Put a cross on the line before the statement that could be more surely proved by using numbers than in any other way; but if there is no such statement, put a cross here........

........Manufacturing is of great financial importance in Germany.
........Chairs are made of wood, nails, screws, glue, cloth, tacks, and paint.
........Swings are made of wood, nails, screws, chains, and other things.
........Mining is an interesting work.
........Iron and steel are in some ways alike and in some ways different.
........Mountains are beautiful.
........Lakes and mountains make beautiful scenery when seen together.
........Islands, lakes, and mountains make beautiful scenery.
........The United States, Canada, and Mexico contain beautiful scenery.
........To the people of the United States the camel is a strange animal.
........The Chinese have customs that are strange to us.
........The Indians have strange customs.

21. Definition

Put a cross on the line before the sentence that tells best what a river is; but if there is no such sentence, put a cross here........

........A river is a body of water; the body of water is in motion; the body of water is sometimes blown upon by the wind.
........A river is flowing water.
........A river is a large stream of water flowing through the land toward some other stream or body of water.
........A river is a stream of water flowing rapidly through the land, and finally reaching the ocean.
........A river is a stream of water flowing in a channel on land toward some other stream or body of water.
........A river is water.

........A river is a large stream of water flowing through a channel toward some other stream or body of water.
........A river is a stream flowing in such a way as to be of use to man.
........A river is a current of water.
........A river is a large current of water.
........A river is a current of water flowing through the land.
........A river is a stream of water flowing through the land; it is used for carrying boats.

22. Recording

Put a cross on the line before the best way of keeping facts in geography until you use them; but if there is no such way, put a cross here . .....

........Remember them even when it is hard to do so.
........Copy them from books.
........Copy sentences from books.
........Keep a note-book.
........Keep several note-books.
........Think hard and carefully as you work.
........Think hard and carefully as you find the facts.
........Let your teacher do it for you.
........Remember everything you read.
........Write each fact briefly and correctly.
........Try to remember the best of what you read.
........Always try to remember the best of what you read.

23. Sufficiency

Put a cross on the line before the way by which you could surely find out how much rain falls in New York City in any one night; but if there is no such way, put a cross here........

........Ask your best friend.
........Ask your teacher at some time when she has not too much to do, so that she will have time to answer you carefully and fully.
........Ask your mother at night when you have reached home after school, at some time when she is not too busy, and when she is not too tired.
........Ask your father.
........Walk one mile alone, so that you may think about it without anyone to talk to you.
........Ask a friend.
........Walk two miles alone, dressed for the weather, so as to be free to think hard about the rain.
........Eat dinner at six, after a long walk in the open air with your best friends.
........Guess at it.
........Ask your father at night, after he has come from work and has finished his dinner.
........Catch rain all night in the right cup rightly placed.
.........Sit in a chair, quietly thinking all the time about the weather and the rain

24. Feasibility

Put a cross on the line before the statement that could be proved true; but if there is no such statement, put a cross here........

........The United States will always be the richest nation in the world.
........The United States will not always be the richest nation in the world.
........The theatre helps educate the world.
........The theatre does not help educate the world.
........China will some day rule the world.
........China will never rule the world.
........All great men are Europeans.
........No great men are Europeans.
........Europe will always be the richest continent.
........Europe will never be the richest continent.
........Asia will sometime give the world the richest men.
........Asia will never give the world the greatest men.

25. Recording

Put a cross on the line before the best way of keeping facts of geography until you are ready to use them; but if there is no such way, put a cross here........

........Remember them by using your own best way of remembering facts.
........Copy them fully from books. Copy the whole of each sentence. Save these copies until you need them.
........Remember where they are so that you may be able to find them again easily whenever you want them.
........Copy sentences from books, leaving out words like "the" and "a," and any other words that have little meaning.
........Keep a note-book in which you write down everything you wish to keep, whether it was found in reading or in observing.
........Keep a note-book. Write each fact.
........Keep each fact you need by writing it in few and chosen words in a note-book.
........Keep a note-book. Use a sentence for each fact.
........Find the true connections among the facts; memorize these, and talk from memory.
........Memorize everything you read.
........Try to remember the important facts you read.
........Always try to memorize the more important facts you read.

26. Clarity

Put a cross before the sentence that is not clear; but if there is no such sentence, put a cross here........

........Some rivers flow into rivers, and some flow into the ocean.
........Many rivers flow into the ocean, but not all flow into the ocean.
........Many brooks flow into lakes, and many brooks flow into rivers.
........Does water always flow down hill?
........Does water ever flow up hill?

........Two plowmen were talking to two other plowmen about their land.
........Flames usually go up, not down or to the side.
........The sun gives us light and heat every day.
........The earth turns around in exactly 24 hours.
........Some years are longer than others, and the longer years are called leap-years.
........Cows give milk, but the milk they give is not all alike; some milk is better than other milk.
........Men milk cows

27. Arrangement

Put a cross on the line before the best arrangement of things; but if there is no such arrangement, put a cross here........

........valley hill mountain; ice steam water; drop brook river
........steam ice water; drop brook river; hill valley mountain
........hill mountain valley; steam ice water; drop brook river
........mountain hill valley; drop brook river; steam ice water
........hill valley mountain; steam ice water; drop brook river
........mountain valley hill; drop brook river; steam ice water
........ice steam water; drop brook river; hill mountain valley
........water steam ice; valley hill mountain; drop brook river
........steam water ice; brook drop river; valley hill mountain
........brook drop river; steam water ice; valley hill mountain
........brook river drop; ice water steam; hill mountain valley
........valley hill mountain; ice water steam; drop brook river
........brook river drop; ice water steam; hill mountain valley

28. Comparison

Put a cross on the line before the sentence that best shows a difference between the Pacific Ocean and the Atlantic Ocean; but if there is no such sentence, put a cross here........

........One ocean is larger than the other.
........The Atlantic Ocean is smaller than the Pacific Ocean.
........The Pacific is much larger than the Atlantic.
........Both oceans have currents of water flowing through them.
........Both oceans are large bodies of water
........The Pacific has more square miles than the Atlantic.
........The Pacific contains many thousands of square miles.
........The Atlantic contains many thousands of square miles.
........Both bodies of water contain salt.
........The Atlantic ocean connects us with far-off lands.
........The Pacific connects us with far-off lands.
........Both oceans are useful, both are large, both carry ships, both have storms, and both are beautiful.

29. Comparison

Put a cross on the line before the best comparison; but if there is no such comparison, put a cross here........

........Lake Huron is larger than Lake Michigan
........Lake Michigan is deeper than Lake Huron.
........Both lakes are very large.
........Lake Michigan and Lake Huron are two lakes in the same water system; they are of nearly the same size; both carry much traffic; both carry many of the same boats. But the former has 22,322 square miles of area and 750 feet of depth, while the latter has only 21,729 square miles of area, but 870 feet of depth.

		Area	Elevation	Greatest Depth
....	Huron	22,322	582	750
	Michigan	21,729	582	870

		Area in square miles	Elevation in feet	Depth in feet
....	Huron	22,322	582	750
	Michigan	21,729	582	870

........These lakes show both likenesses and differences. The following are likenesses: (1) Both are in the same water-system. (2) Both carry much of the same traffic. (3) They have the same elevation, 582 feet. (4) Their areas are nearly equal, Huron having 593 square miles more than Michigan. (5) The lakes are much alike in shape, Huron being about three times as long as wide, and Michigan about four.
........These lakes show both likenesses and differences. The following are likenesses: (1) Both are in the same water-system. (2) Both carry much the same same traffic. (3) They have the same elevation above sea-level, 582 feet. The following are differences: (1) Huron has 22,322 square miles of area, and Michigan 21,729. (2) The lakes differ in shape, Huron being about three times as long as wide, Michigan about four. (3) The greatest depth of Huron is 750 feet, Michigan 870.
........Huron and Michigan are nearly the same size.
........Huron is shorter than Michigan.
........Lake Huron does not have Chicago upon its shore.
........Both lakes are useful.

30. Classification

Put a cross on the line before the best grouping of things; but if there is no such grouping, put a cross here........

	cat horse crab man fish dog		cat horse crab man fish dog
....	tree bush wood man apple	....	tree bush wood apple fruit
	stone iron lead tin steel		stone iron man lead tin

cat horse crab man fish dog
........tree bush wood apple fruit
stone iron lead tin steel

cat horse crab man fish dog
........tree bush wood apple fruit
stone iron bush tin steel

cat horse crab man fish dog
........tree bush man wood apple
stone iron lead tin steel

cat horse crab man fish dog
........tree bush tin wood apple
stone iron lead tin steel

cat horse crab man tin dog
........tree bush wood apple fruit
stone iron lead tin steel

cat horse crab man dog
........tree bush wood apple fruit
stone iron lead tin man

cat horse crab man fish dog
........tree bush wood apple fruit
stone iron lead man steel

cat horse crab man fish dog
........tree bush iron wood fruit
stone iron lead tin steel

cat horse crab man fish dog
........tree bush stone apple fruit
stone iron lead tin steel

cat horse crab man fish fruit
........tree bush wood apple fruit
stone iron lead tin steel

31. Relevancy

Put a cross on the line before the fact that helps show that air is colder high up than it is low down; but if there is no such fact, put a cross here........

........More snow fell in the spring of 1888 in New England than for many springs before.
........Northern people moving to the south can live best in the high regions.
........Men who fly in flying machines must sometimes dress warm.
........It is colder in a well than at the top of it.
........It is much colder in the northern part of Alaska than in the southern.
........High towers have been blown down by the wind in many parts of the world.
........The Washington Monument is 555 feet high.
........Trees will not grow at the north pole, nor will they grow at the south pole.
........Bears sleep through the winter.
........Many birds that live in the north during the summer fly south for the winter.
........A mountainous country is more beautiful than a flat one.
........It often feels colder out of doors than in a room with windows and doors shut.

32. Statistics

Put a cross on the line before the statement that would need the most numbers to prove it; but if there is no such statement, put a cross here........

........Many plants grow fast in warm, wet places.
........Many men raise wheat in the uppper Mississippi valley.
........Wheat is the chief crop of the upper Mississippi valley.

........Canada raises more wheat and corn than the United States.
........Large quantities of corn and oats grow in the United States.
........The United States is a great manufacturing nation.
........Water covers two-thirds of the earth's surface.
........More rain falls on the western coast of this country than upon the eastern.
........China makes 15 million pounds of silk a year, Japan 7 million.
........It is sometimes cold in Canada.
........It is sometimes cold in England.
........It is sometimes cold in the United States.

33. Sufficiency

Put a cross on the line before the way in which you could surely find out how much rain falls in any one year in New York; but if there is no such way, put a cross here........

........Save the newspaper reports.
........Ask a good weather-man.
........Catch rain in a dish every time it rains; measure the amount; save the records.
........Take the government statement of the average rainfall per year.
........Watch the rains yourself, carefully guessing the amount each time it rains, and saving the records.
........Catch rain all the time it rains in a dish carefully made for that purpose.
........Read rainfall reports from all the important cities of the United States.
........Take the daily weather report of the newspapers, save the figures, and add up at the end of the year.
........Read newspapers and reports.
........Ask a good teacher in geography.
........Read all the New York papers.
........Discuss the matter carefully with a group of friends.

The Right Answers

Each exercise is either right or wrong. Mark wrong all cases of crosses on more than one line. Mark wrong all cases of really ambiguous placing of a cross. Any mark which clearly indicates an answer here designated as a correct answer, should be marked correct. Some children mark with a slanting or vertical line. This should be counted correct. The right answers are below. No other answers may be counted correct.

1 x dog
2 x John
3 x Mountain
4 x A small
5 x The
6 x baby, boy, young
7 x it
8 x Write the facts
9 x Look

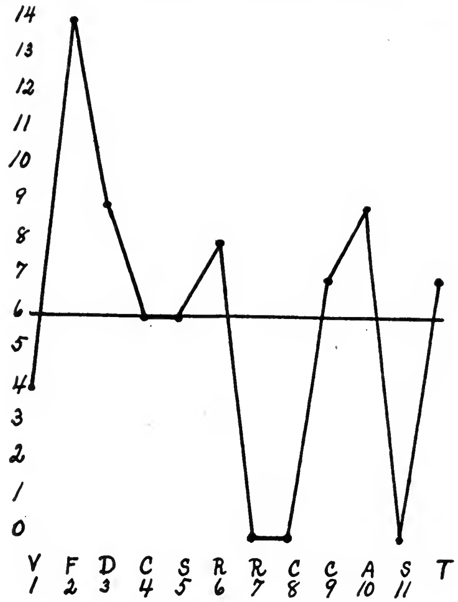
10 x cloud, rain, brook
11 x Will
12 x A small body of water
13 x Canada
14 x There is snow
15 x China
16 x A pound
17 x house school-house
18 x Will
19 x Will
20 x Manufacturing
21 x A river is a large stream of water flowing through the land
22 x Write
23 x Catch
24 x The theatre helps
25 x Keep each fact
26 x Two
27 x valley hill mountain; ice water
28 x The Pacific is
29 x These lakes . . . . Michigan 870.
30 x (The second group at the left.)
31 x Northern .
32 x The United
33 x Ask a good weather man.

MEANS.—The means is found in the analytical diagnosis of classes and of individuals afforded by the test. This is illustrated in the case of one class and of one pupil below:

This illustration suggests that L. S. needs improvement in the abilities here named recording, comparison, and sufficiency. She seems especially good in her judgment of the feasibility of problems, and about average in her other abilities. The individual graph is, of course, less reliable than the grade graph would be. Suppose, however, that the ability of L. S. to tell good from bad comparisons were really near the median of her class instead of far below it, as indicated by her graph. If this fact is discovered, the graph has served its purpose. Taken with the extended observations of a teacher, and perhaps of a principal, the graph will constitute a real value, will offer a tentative analysis of the child's ability, and will suggest means of improvement.

At the outset two courses were open: to study one ability intensively, or to survey in a preliminary way a field made up of several abilities; to investigate, for instance, judgments as to the values of different problems (ability 1), or to look more broadly and less

SCHOOL A; PUPIL L. S.; AGE 11; CLASS 6A; DATE, MARCH 14, 1917



Graph showing the performance of a sixth grade class and of individual L. S. The vertical scale at the left is of P. E. units of difficulty. The horizontal row of numbers and letters is read thus: Ability 1, value; ability 2, feasibility; etc. (See page 5.) The variations of L. S. above and below her class appear. T means time.

deeply upon judgments involved in the exercise of the eleven abilities. The more inclusive scope was selected, the test in the present form being the result. Through this test conclusions as to the group standing of a class of pupils in each of the eleven abilities may be drawn. This preliminary survey of portions of the great field of scientific ability may properly be followed by a more searching one composed of a series of tests, one or more for each of the eleven abilities and for other abilities. Through such a series conclusions could be drawn regarding the abilities of individual pupils,—conclusions which would possess a degree of reliability not contemplated in the present work.

RELIABILITY.—The reliability of conclusions expressed, upon the basis of the present test, by graphs for classes of pupils, would be sufficient; for individuals, usefully suggestive. Consider, however, the case for the reliability of the individual graph. Exercise number one calls for at least twelve classification judgments before the right answer can be made with confidence. Since there are

thirteen possible places for a cross, the chance for getting an exercise right without thinking is about eight per cent. The probability, therefore, that a cross on the line before the word *dog* is placed there as the result of careful thought is great, while the probability that three crosses on the right lines in the set of exercises labelled *classification* are placed there as the result of careful thought is very great. Students getting all three right really know something about classification. There would be of course a negligible number of exceptions in a large number of cases. It is true that an exception is not negligible for the individual that happens to be concerned, but the frequency would be low with which more than one or two of the ability scores of an individual would be much wrong. This graph is a usefully suggestive conclusion. It is at the same time the most vital and the least reliable conclusion. It will be important to develop at a later time two commensurate series of scaled exercises for each ability, in order to obtain very accurate graphs for each child.

A precise measurement of one trait in one individual by means of the present test is not contemplated, for such a measurement could not be based upon responses to three exercises. Four scores only are possible: none right, one right, two right, and three right.

CLASSIFICATION OF THE TEST.—This is a test of method, not of content. One class of mental measurements concerns the subjects of the curriculum, as spelling, reading, writing, and arithmetic. It deals with the content of subjects. Another class studies capacities and abilities each of which is used in several or in all of the subjects of the curriculum. In this class belong measurements of intelligence, memory, and reasoning. In this class also belongs this scientific method test, which aims to measure the abilities of students to use certain elements of the scientific method. All the content of this test falls within the subject of geography, but the aim is to measure the ability to use scientific method in geography, not the ability to remember geographical content.

Both method and content are of worth. To whatever extent an ability to employ a method may be transferred from field to field by students,—and even if it may not be transferred at all, it is still an essential to progress within the field. Content is the other essential. It is the whole world. Its study broadens the view. It is the material upon which we use method. Yet it means relatively little to our matriculating boys and girls without some conscious mastery of method.

The History of the Test.—The test is in its third edition. After an experiment with a brief first edition on October 31, 1916, all the exercises were for one reason or another discarded. It had consisted of nine exercises somewhat similar in form to those of the present edition. A second edition of forty exercises was then prepared. There were three or four exercises to measure each of the eleven abilities. These exercises were submitted to six judges for two purposes: to determine the validity of the answers and the fitness of each exercise as a measure of a specified ability. For instance, numbers 3, 14, and 31 are intended to measure relevancy judgments. Do they do this, or do they do something else? Are the students who score high in these three exercises the ones whose relevancy judgments are usually good? The question is a fair one. Within limitations there is an answer. The six judges were chosen for mental acumen and for their long study of scientific method. One is the head of a department of philosophy, of which two others are members; one is a member of a department of history, and one of a department of mathematics,—all these in a university; and one is an engineer and a teacher of mechanical drawing in a high school. Each spent from two to five hours upon the exercises. They worked separately, and except in one case had not seen the exercises before doing the work. To each were presented the forty exercises and the list of abilities substantially as defined above. Each stated which, if any, of the abilities was tested by each exercise. Not even at first was there a discouraging amount of disagreement. Some exercises were discarded, some modified slightly, and some rebuilt. At last the judges came to a practical agreement. Below is a tabulation of the responses of the judges, (a, b, c, d, e, f,) to the question: Does each exercise test that which it is designed to test? As to the correctness of the answers, each judge passed upon that point for each exercise as a preliminary to the tabulated judgment above. Each exercise needing it was modified until it met the approval of each judge before whom it came. The v's serve therefore as a tabulation of this judgment also. Every exercise has only v's because each one not at first receiving them was modified until it did.

	a	b	c	d	e	f
1	v	v	v	v	v	
2	v	v	v	v	v	
3	v	v	v	v	v	
4	v	v	v	v		
5	v	v	v	v	v	
6	*v*		*v*	*v*	v	v
7	v	v	v	v	v	
8	v	v	v	v	v	
9	*v*	v		v	v	v
10	*v*	*v*	v	v	v	v
11	*v*	v	*v*	v	v	
12	v	v	v	v	v	
13	v	v	v	v	v	
14	v	v	v	v	v	
15	v	v	v	v	v	
16	v	(v)	v	v	v	
17	v	v	v	v	(v)	v
18	*v*	v	*v*	v	v	
19	*v*	v	*v*	v	v	
20	v	v	v	v	v	
21	v	v	v	v		
22	v	v	v	v		
23	*v*	v	v	v	v	v
24	v	v	v	v		
25	v	v	v	v	v	
26	v	v	v	v	v	
27	(v)	v	v	v	v	
28	v	v	v	v	v	
29	v	v	v	*v*		v
30	v	v	v	v	(v)	v
31	v	v	(v)	v		
32	v	v	v	v		
33	*v*	*v*		v	v	v

v......Judge gave the opinion that the exercise would test the use of the ability with which it is labelled.

v......Judge gave such an opinion after some modification of the exercise, which was then submitted to the other judges indicated.

(v)......Judge agreed with the label after a brief discussion.

The exercises have also been through the processes of renovation and rebuilding as the result of two trials, one with the first and one with the second edition. It would be too much to state that each set of three exercises measures all that which is defined under its catch-word and nothing more; but it is hoped that each set of three is a significant measure of some considerable portion of that which is defined and of little else.

In March, 1917, the forty exercises of the second edition were submitted to 140 grade school pupils and to 141 high school pupils. So far as could be seen most of the exercises were good ones. Seven were discarded, leaving three for each of the eleven abilities,—thirty-three in all. The text of the test seemed now to have been fitted for a more extended use. About 1800 grade pupils took the test in May, 1917, and others will soon take it in grade schools, high schools, and colleges.

Throughout the building of the test, it has been the aim to reduce toward a minimum the difficulties of the directions found at the beginning of each exercise, of the wording, and of the content of

the test. Yet the trite, but pertinent question: What is measured? is sure to make its appearance. The safe answer is: Responses to the thirty-three exercises of the test, as printed, are measured. The question: Does the test measure abilities as defined? must, in the last analysis, be answered by those who read the statistics and conclusions to be published at some future time.

Conclusion.—Our schools now confront a danger arising from an embarrassment of riches. They are to be felicitated on having in their stock-rooms splendid supplies of measuring tests, which are being put to excellent use,—tests in the four fundamental processes and the reasoning of arithmetic, in punctuation, in spelling, in language, in oral and silent reading, in handwriting, and in drawing. But it happens to be generally true that those human abilities which it seemed proper to measure first, perhaps because they could be measured easily, did not happen to be those of the greatest value socially. Such measurements as those of ability in silent and rapid reading, of intelligence, of thinking, and of judgments are of the utmost value. Coupled with the tendency to measure first that which seemed most feasible, there has been, fortunately, some tendency to measure first that which seemed most important,—provided it seemed also feasible. The latter tendency is right. Together with the further development of tests which involve mainly memory, we need a proportionate production of tests of the highest mental qualities,—so far as these may be measured. While the measurement of scientific ability is beset with difficulty, still it is dealing, not mainly with memory, nor addition, nor spelling, nor punctuation,—important as these are,—but with the noblest intellectual function, judgment.

A STANDARDIZATION OF CERTAIN OPPOSITES TESTS

HARRY A. GREENE, A. M.

State University of Iowa

The frequency of the use of certain forms of opposites tests in educational measurements, and the unusual reliability of the results make a careful standardization of these tests desirable. Early results, such as those obtained by Simpson,* and by Bonser† are certainly significant; but the lack of uniformity in the type of stimulus, in methods of giving the tests, and in determining the correct response words, have thus far made comparison of results obtained almost impossible.

One of the first attempts‡ at the standardization of tests of this type was made by Professor Irving King and Hugo Gold¶ of the State University of Iowa in 1915–16. "The object of this study was to take certain lists of words which have been used by various experimenters for opposites tests, to give these tests under the best possible conditions, and to attempt to standardize a list or lists from which one might expect reasonably uniform results (when given under like conditions); and furthermore to standardize the grading of opposites tests by determining the words likely to be given as the opposite of each word, their frequency, and their relative values." Four lists of twenty easy stimulus words, and four lists of twenty hard stimulus words as used by Simpson were selected. These lists were given individually to one hundred subjects consisting of nine faculty members, twenty-three graduate students, forty-seven seniors, and twenty-one juniors, all in the departments of education and psychology in the University of Iowa.

On the basis of these responses the following graded list of Simpson's Hard Stimulus words was recommended. The order of the words indicates the difficulty beginning with the easiest:

---

*Simpson—*Correlation of Mental Abilities;* Teachers College, New York, 1912.

†Bonser—*Reasoning Ability of School Children;* Teachers College.

‡An earlier standardization of a different list was made by Woodworth and Wells.

¶King and Gold—*A Tentative Standardization of Certain Opposites Tests;* Journal of Educational Psychology, Vol. VII, 459–482, 1916

Remember, after, sell, wild, straight, to win, beautiful, increase, rough, succeed, strength, silly, to hold, generous, to bless, over, simple, part, refined, despondent, to take, to spend, to lack, to float, vertical, apart, belief, stingy, active, to degrade, cruel, broken, innocent, to reveal,

busy, permanent, permit, sleepy, sure, respect, preserve, miser, rude, frequently, tardy, level, stormy, conservative, strict, ignorant, deceitful, exciting, motion, sinful, weary, rigid, genuine, proficient, impoverish, grand, result, imaginary, suspicious, precise, pride, unless, suave, forcible.

In the early use of opposites tests various methods of scoring were used. In some cases a time limit was set, and the scores calculated in terms of the number of stimulus words to which the correct responses were made. Other investigators, objecting perhaps to placing a time limit on the test, preferred to score the results in terms of the time required for the subject to complete the test, plus a penalty for each error or omission. In both of these methods no account was taken of the wide differences in the difficulty of the various stimulus words. Each error or omission was regarded as of the same value and the person tested was penalized the same amount regardless of the word missed. Clearly, if the best results are to be realized from the tests, the relative values of the different words must be taken into consideration in grading the responses. The previous study, while standardizing the *responses*, did not definitely evaluate the difficulty of the stimulus words in such a way as to make them readily usable in scoring results. With this in mind the writer attempted to secure such an evaluation for each stimulus word given in the list above,—this evaluation to be based on a large number of cases, and to be in terms of the difficulty of the word as indicated by the character and the accuracy of the responses given to it.

There are in general two methods that may be used in the determination of the values of these stimulus words. By the first method the values are determined in terms of the time required for the subject to respond to the stimulus word correctly. By the second the evaluation is determined on the basis of the percentage of accuracy of response for each word. It is quite likely that there would be a high correlation between the two procedures. That is, the words, that would require a long time on the average to elicit the correct response would in the long run be the ones that would

stand lowest in percentage of accuracy. The first method seems to offer certain possibilities where the subjects are to be tested individually. However, the gratifying results obtained from using the test as a group test, lead one to think that an evaluation based on the actual usage of the test would be more reliable. Accordingly, in the present study the second method was followed.

After having decided upon the method of evaluation the writer was then confronted with two further possibilities. The evaluation could be made directly in terms of the percentage of failures, or could be turned into percentile values by using the normal or probability curve as a basis. For the sake of comparison both procedures were carried through, although the results from the second method only will be presented here. The chief differences noted were that the easy words received a higher point value on the percentile basis, and the hard words a correspondingly lower value. For example, the first five words in Table I which are given a value .5 and .6 points were evaluated at .1 points by the first method, and *suave*, which is evaluated at 2.0 points, was given 2.4 points by the first method. The actual rank correlation between the results of the methods was almost perfect, "r" being equal to .996.

The stimulus words listed above were given in two groups to freshmen in the University of Iowa in the fall of 1916 and in 1917.* As a result of these investigations, responses from 990 individuals to the first half of the list, and 710 responses to the second half of the list were available. Using the King and Gold standardization as the key to the correct responses, these papers were corrected, and the responses to each stimulus word were tabulated.† In this tabulation the words given as *full value* were counted as *one*, and the *half value* words were counted as *one-half*. For example, if it was found that in responding to the stimulus word *to win*, 700 persons gave *to lose* (full value), and 72 gave *to fail* (half value), the total number of responses was counted as 736. In this way the totals were found for each of the eighty words in the list. Since these totals represent the number of correct responses, it is apparent

*In 1916 by James L. McCrory, and in 1917 by the writer, both under the direction of Prof. Irving King.

†Words given 2–3 value in the King and Gold table were given half value in this study, and all words counted as 1–3 by them were counted as wrong. This was mainly to save time by simplifying the procedure.

that the relative difficulty of the stimulus words is represented by the complement of these totals. That is to say, a word with a small total is more difficult than a word with a larger total. To secure the values on this basis the total number of correct and half correct responses were subtracted from the total number of possible responses, in this case 990. In the case of the words from the last half of the list for which there were only 710 responses available, the totals were increased in proportion as 990 is to 710, thus making all the words in the list comparable in spite of the differences in the number of possible responses.

These totals were turned into percentage of failure by dividing each by 990, the total possible responses. The percentage scores were changed directly into percentile values by reading directly from table based upon the area of the probability curve, assuming the base line to be broken off arbitrarily at $+3\sigma$. The particular abilities involved in this test are assumed to fit the probability curve, and the percentage of subjects who answer the various stimulus words correctly correspond to percentages of area under the curve from the 0 point to a point on the base line. This point on the base line in units of sigma was transformed into percentile values by setting 0 at -3σ, 50 at the mean, and 100 at $+3\sigma$. To illustrate the use of the table let us locate a word in the percentile scale. For example, the word *suave* was missed by 96% of the subjects. From the table* a word missed by 96% of the cases is valued at 79. In this manner the percentile values were found for the 80 words.

These values were then totaled, and each value divided by the total, thus turning the percentile values into relative point values. These values range from .5 point for the words *beautiful* and *straight*, up to 2.0 points for the word *suave.* (See Table I.) It is evident that if the words were of uniform difficulty the point value of each word would be 1.25. Table I shows the words of the same point value grouped together in ascending order of difficulty. By making use of this table it is possible for one to compare directly the difficulty of the various stimulus words; or to select a number of tests of the same difficulty and uniformity. By using the complete list of eighty words and considering each word at its point value a test with a total possible score of 100 points is obtained. This is valuable where it is desirable to combine directly the results of the two opposites tests into one score.

*Rugg, H. O.—*Statistical Methods Applied to Education;* Houghton, Mifflin Co., Chicago, 1917. Table VI; pg. 396.

TABLE I

List of Opposites Stimulus Words Arranged in Ascending Order of Difficulty with Point Values

Point Value	Stimulus words:	Point Value:	Stimulus words:
.5	beautiful	1.3	level
	straight		over
.6	after		refined
	increase		respect
	remember		rude
.7	rough		sleepy
	stingy		strict
.8	apart		to reveal
	sell	1.4	active
	to float		frequently
	take		imaginary
.9	generous		injurious
	strength		permit
	succeed		rigid
	to lack		to degrade
	vertical	1.5	conservative
	wild		ignorant
1.0	cruel		past
	innocent		permanent
	part		sinful
	simple		weary
	to bless	1.6	exciting
	to win		deceitful
1.1	busy		genuine
	hindrance		grand
	insignificant		haughty
	miser		motion
	preserve		proficient
	sure		result
	to hold		suspicious
1.2	belief		venturesome
	broken	1.7	animated
	despondent		disastrous
	diligent		impoverish
	silly	1.8	forcible
	stormy		precise
	tardy		pride
	tender		unless
	to spend	1.9	serious
1.3	clumsy	2.0	suave

It was of some interest to the writer to see the extent to which the standardization on the basis of the 990 responses was more reliable than the earlier standardization on the basis of the one hundred responses. A correlation was made between the list of eighty words as evaluated by King and Gold and by the present study. Using the Pearson method adapted to rank differences an "r" of .895 was found. The correlation between the two methods of evaluation discussed earlier in this paper was given above.

The stimulus words with their point values, full value responses, and half value responses, are shown in alphabetical order in the following pages for easy reference. The number preceding the stimulus word in each case is the *point value:* the stimulus word appears in small capitals, the full value responses in italics, and the half value responses in ordinary type.

1.4 ACTIVE: *passive, inactive;* slow, sluggish, lazy, quiet.
.6 AFTER: *before.*
1.7 ANIMATED: *lifeless;* spiritless, quiet, dead, sluggish, unanimated.
.8 APART: *together, with, joined;* near, adjacent.
.5 BEAUTIFUL: *ugly, homely.*
1.2 BELIEF: *disbelief, doubt, unbelief, distrust.*
1.2 BROKEN: *whole, smooth, mended, connected, continuous, intact;* continued, level, repaired, unbroken.
1.1 BUSY: *idle;* lazy, indolent, unemployed, unoccupied.
1.3 CLUMSY: *graceful;* agile, nimble, skillful, dextrous.
1.5 CONSERVATIVE: *radical, progressive, liberal;* hasty, reckless.
1.0 CRUEL: *kind, humane;* tender, gentle, kindly.
1.6 DECEITFUL: *frank, straightforward, sincere, ingenuous;* truthful, honest.
1.2 DESPONDENT: *happy, cheerful, hopeful, courageous;* gay, joyful, joyous, light-hearted.
1.2 DILIGENT: *lazy, slothful;* idle, careless, negligent, neglectful.
1.7 DISASTROUS: *fortunate, lucky, auspicious;* beneficial, helpful.
1.6 EXCITING: *quiet, quieting, soothing;* calming, monotonous, uninteresting.
1.8 FORCIBLE: *ineffective, powerless;* weak, gentle, mollifying.
1.4 FREQUENTLY: *seldom, rarely, infrequently;* occasionally, usually, almost-never.
.9 GENEROUS: *stingy, selfish;* miserly.
1.6 GENUINE: *spurious, counterfeit;* false, artificial, fake.
1.6 GRAND: *mean, simple, lowly, humble;* small, plain, poor, low, little, common, insignificant.
1.6 HAUGHTY: *humble;* low, lowly.
1.1 HINDRANCE: *help, aid, assistance, benefit;* advantage.
1.5 IGNORANT: *learned, knowing, informed;* wise, intelligent, well informed.
1.4 IMAGINARY: *real;* realistic.
1.7 IMPOVERISH: *enrich;* replenish.
.6 INCREASE: *decrease, diminish, lessen.*

1.4 INJURIOUS: *helpful, beneficial;* healthful, safe.
1.0 INNOCENT: *guilty;* wise, knowing, sinful.
1.1 INSIGNIFICANT: *important, significant, prominent;* great, grand, distinguished.
1.3 LEVEL: *rough, hilly, undulating, mountainous;* uneven, broken.
1.1 MISER: *spendthrift;* philanthropist, benefactor.
1.6 MOTION: *rest, repose;* quiet, inaction, inactivity, silence, quietness, at rest, quietude. quiescence. motionless.
1.3 OVER: *under, underneath;* below.
1.5 PAST: *future, coming.*
1.0 PART: *whole, unit, all, connect.*
1.5 PERMANENT: *transitory, transient, ephemeral, fleeting;* temporary, temporal, changing, changeable.
1.4 PERMIT: *refuse, forbid, hinder;* deny.
1.8 PRECISE: *inexact, inaccurate.*
1.1 PRESERVE: *destroy, spoil, ruin, waste, squander.*
1.8 PRIDE: *humility, self-abasement;* humiliation, humbleness.
1.6 PROFICIENT: *inefficient, incompetent;* deficient, incapable, unskilled.
1.3 REFINED: *coarse, rude, uncultured, crude, unrefined;* uncouth, uncultivated.
.6 REMEMBER: *forget.*
1.3 RESPECT: *despise, disrespect, dishonor, disregard, disdain;* scorn, abhor.
1.6 RESULT: *cause;* beginning.
1.4 RIGID: *lax, loose;* flexible, limp, relaxed, flabby, lenient, elastic, limber, pliable.
.7 ROUGH: *smooth, polished, plain.*
1.3 RUDE: *polite, courteous, refined, genteel;* cultured, gentle, cultivated, urbane, well-bred.
.8 SELL: *buy, purchase.*
1.9 SERIOUS: *frivolous;* gay.
1.2 SILLY: *wise, serious, sensible;* sober, sane, dignified.
1.0 SIMPLE: *complex, complicated, elaborate, intricate;* hard, difficult.
1.5 SINFUL: *sinless, holy, godly;* good, righteous, pure, innocent. virtuous.
1.3 SLEEPY: *wakeful;* awake, wide-awake, fresh.
.7 STINGY: *generous, liberal, lavish, free-hearted;* free, extravagant, spendthrift.
1.2 STORMY: *calm;* clear, peaceful, quiet, fair, still.
.5 STRAIGHT: *crooked, curved, bent.*
.9 STRENGTH: *weakness;* weak.
1.3 STRICT: *lax, loose, lenient. liberal;* careless.
1.1 SURE: *uncertain, doubtful, undecided.*
2.0 SUAVE: *impolite, brusque;* rough, rude, coarse, uncouth.
.9 SUCCEED: *fail.*
1.6 SUSPICIOUS: *trusting, trustful, unsuspicious, unsuspecting;* innocent, credulous, confiding.
1.2 TARDY: *punctual, prompt;* early, on-time, before-time.
1.2 TENDER: *tough, harsh;* hard, cruel, rough, hard-hearted.
1.0 TO BLESS: *to curse, to condemn, to execrate.*
1.4 TO DEGRADE: *to elevate, to uplift, to exhalt*
.8 TO FLOAT: *to sink.*

1.1 TO HOLD: *to let go, to drop, to let loose, to release, to loose, to loosen, to let fall.*
.9 TO LACK: *to have, to possess, to be supplied;* to abound.
1.3 TO REVEAL: *to hide, to conceal, to secrete;* to obscure.
1.2 TO SPEND: *to save;* to earn, to get.
.8 TO TAKE: *to give, to leave, to let go, to loose, to refuse.*
1.0 TO WIN: *to lose;* to fail.
1.8 UNLESS: *if.*
.9 VERTICAL: *horizontal.*
1.6 VENTURESOME: *careful, cautious;* cowardly, timid, afraid, fearful.
1.5 WEARY: *rested, refreshed;* fresh, refresh, rest.
.9 WILD: *tame, domesticated, civilized;* cultivated.

100.0 points

THE DIAGNOSTIC VALUE OF THE WOODY ARITHMETIC SCALES: A REPLY.

PART II

W. W. THEISEN AND CECILE WHITE FLEMMING

Bureau of Educational Measurements, State Department of Public Information, Madison, Wisconsin

From a constructive point of view the criticism that the Woody tests do not diagnose may be answered by giving the results of typical analyses of performances for a number of classes.

1. A sixth grade of 33 pupils tested in subtraction.

The examples of the Woody Subtraction Scale, Series A, are shown below.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
8	6	2	9	4	11	13	59	78	7−4=	76
5	0	1	3	4	7	8	12	37		60

(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
27	16	50	21	270	393	1000	567482	$2\frac{3}{4}-1=$
3	9	25	9	190	178	537	106493	

(21)	(22)	(23)	(24)	(25)	(26)
10.00	$3\frac{1}{2}-\frac{1}{2}=$	80836465	$8\frac{7}{8}$	27	4 yds. 1 ft. 6 in.
3.49		49178036	$5\frac{3}{4}$	$12\frac{5}{8}$	2 yds. 2 ft. 3 in.

(27)	(28)	(29)	(30)
5 yds. 1 ft. 4 in.	10−6.25=	$75\frac{3}{4}$	9.8063−9.019=
2 yds. 2 ft. 8 in.		$52\frac{1}{4}$	

(31)	(32)	(33)	(34)	(35)
7.3−3.00081=	1912 6 mo. 8 da.	$\frac{5}{12}-\frac{2}{10}=$	$6\frac{1}{8}$	$3\frac{7}{8}-1\frac{5}{8}=$
	1910 7 mo. 15 da.		$2\frac{7}{8}$	

The median score for the class on this test was 21.9, or more than three examples below standard. The type and frequency of error made by this class of pupils are shown in Table IV. The number of examples omitted by each pupil and the number of times any given example was omitted are also shown.

TABLE IV

Showing Number of Each Type of Error Made by Each of 33 Sixth Grade Pupils on the Woody Subtraction Test, Series A

	Pupils																																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	Total
1. Incorrect subtraction of denominate numbers	2	1	1	2	3	3	1	1		3	3	3	3	2		3	3	3		3	3	3	3	3	2	3	3	3	3		3	3	3	75
2. Errors in subtraction of decimals due to incorrect placing when operation is indicated by minus sign	1	2		2	3	3	3			1	1	2	2	2			3	3		3	2	3	2	3	1	3	3	2	3		3	1	2	59
3. Incorrect subtraction of fractions when both minuend and subtrahend contain fractions	1		3		1	2	1	1	2	1	1	2	1			2	1	2		2	4	1	4	2	2	2		3	1	1	3		4	50
4. Not reducing fractional answer	3	3		3	3		3		1		3		1	2	2	1	1	3		2	1	2	2	1	1		3	3			1	3	2	50
5. Wrong operation															26	1	4					2							4					37
6. Subtracting as if not borrowed				3								3					8		2	3		1		2	1			3	1		3			30
7. Subtracting as if borrowed	3			1	2						1			3				3		2		4	5						3	2				29
8. Error in fundamental combinations			2		1		1					2		3		2		1		2	2		3	1			1	1	1			1	2	26
9. Omission of, or incorrect placing of decimal point	2			2	3		2			1	1	1	1	1			1	1		1			1				2		3		1		2	26
10. Incorrect subtraction of fraction from whole number or vice versa	1	1	1		3	1	1		1	1			1	1		1	1	1		1	1	1	1				1	1	1		1	1	1	25
11. Failure to recognize the sign of operation	1	1	3	1		1																										1		8
12. Failure to classify result	1																																	1
Number of examples omitted	1		4					1	2	3				4	5	4			6			1			1				1	6		2	1	42

Note: The examples omitted were No. 24 four times, 25 once, 26 twice, 27 twice, 28 five times. 30 seven times, 31 nine times, 32 three times, 33 seven times, 34 and 35 each once.

The greatest number of errors, 75 out of a total of 99 chances, was made in examples in subtraction of denominate numbers. Apparently this was due to the omission of this subject matter from the course of study. Reference to the original tabulation sheet indicating scores by pupils for each example shows that two pupils solved all three of the examples in subtraction of denominate numbers. Two others had two of the three correct. When the subject of denominate numbers is presented as a topic to this class, shall these pupils who apparently already possess the ability to subtract denominate numbers be asked to spend the same amount of time on the subject as do the other 29 who have need of this instruction? Has our analysis thus far revealed anything as to the needs of different pupils in this class? Has the test helped us to diagnose to the extent of revealing the presence of certain abilities in certain pupils of the class?

The second most common source of error occurred in examples 28, 30, 31. These are:

$10-6.25=$, $9.8063-9.019=$, and $7.3-3.00081=$

In each of these the operation is indicated by the minus sign. Fifty-nine errors out of a possible ninety-nine were made in these examples because of incorrect placing of the subtrahend with respect to the minuend in performing the subtraction. Twenty-eight pupils contributed errors of this sort. In answer to Example 28, 6.15 was common, obtained by disregarding denominations, placing the 10 under the .25 and then subtracting. In example 30, the answer 8.9044 often occurred. This is obtained through a manipulation such as this

$$\begin{array}{r} 9.8063 \\ 9.019 \\ \hline 8.9044 \end{array}$$

3.00008 occurred several times in answer to $7.3-3.00081=$. This remainder is obtained by placing the 7.3 directly under the 81 of the subtrahend which is then treated as the minuend. The recurring incorrect answers to these examples due to the same cause were several, depending on where the pupil chose to place the respective figures.

Fifty errors occurred through inability to subtract correctly numbers with fractions (either mixed or pure fractions) and an equal number of errors through failure to reduce to lowest terms, or through errors in reduction. The latter type is made up of ex-

amples in the subtraction of fractions in which the operation was correctly performed, but in which answers either were not reduced (*e.g.*, $2^2/_8$ occurs often in answer to example 35, which is $3\frac{7}{8}-1\frac{7}{8}=$) or were incorrectly expressed as $2^2/_8=\frac{1}{4}$. Thus fifty errors, or an average of more than one per pupil, were made by this class by reason of the fact that the teacher had neglected to teach the class to express answers in lowest terms. Only nine of the class did not make an error of this sort. Ten made three errors each.

Twenty-five errors were due to difficulties in subtracting a fraction from a whole number or vice-versa, *e.g.*, in examples 20 and 25, $2\frac{3}{4}-1=$ and $\begin{array}{r}27\\ \underline{12\frac{5}{8}}\end{array}$. Example 20 was missed in all six times, and example 25 twenty-four times. It is significant that for example 25 the answer was given as 12 $\frac{5}{8}$ by fourteen of the class. Another gave $15^0/_8$. The recurrence of the same error is sufficient to indicate a type of subtraction in which the class has not yet attained proficiency.

What seem to be two common and somewhat similar sources of error appear in the subtraction of whole numbers. In one the pupil treats a figure of the minuend as if he had previously borrowed from it when he really has not. In the other he proceeds as if he had not borrowed when as a matter of fact he has. This type of error occurs most frequently in examples 19 and 23:

$$\begin{array}{r}567482\\ \underline{106493}\end{array} \text{ and } \begin{array}{r}80836465\\ \underline{49178036}\end{array}.$$

Twenty-nine of the former and thirty of the latter were found. It is granted, of course, that this classification may include pure errors in combination. In both types errors were limited to eleven pupils. Of the type in which pupils subtracted as if they had not borrowed, eight errors were made by one pupil. Five others made three each. These pupils at any rate reveal a need of drill in subtraction of whole numbers with and without borrowing. It is an interesting fact that seven of the eleven pupils who made errors through subtracting as if they *had* borrowed made no errors of subtracting as if they *had not* borrowed. Seven of the eleven who made the latter type of error did not make the former. These facts are at least highly suggestive of individual needs in the way of drill for the removal of specific weaknesses in subtraction. How many more

errors does the pupil who made eight errors of the first of these types need to make before it can be said that we have located one of his specific needs?

Omitting the decimal point or incorrectly placing it occurred in examples 21, 28, 30 and 31. Twenty-six errors of this type were contributed by seventeen pupils. Two pupils made three errors each of this type. Little is indicated as to the specific needs in subtraction of pupil 15 who contributed 26 of the 37 cases listed under "wrong operations" and omitted 5 other examples. However, the teacher may well observe the habits which this pupil and also pupils 17 and 29 are forming toward following specific directions. Individual attention to the physical condition and mental capacity of these children and further testing would result in a more adequate diagnosis of their difficulties.

Thus we have diagnosed marked weaknesses of this class to the extent of discovering: (1) The inability of this class to add denominate numbers; (2) a habit of disregarding denominations in subtracting decimals when the sign of operation is expressed by a minus sign; (3) weaknesses in subtraction of fractions—fractions from fractions, and fractions from whole numbers or vice-versa; (4) failure to express answers in lowest terms; (5) weakness in placing the decimal point; and (6) errors in borrowing.

2. A fifth grade of 41 pupils tested in subtraction.

The median score for this class was 21.2 or .8 above the Woody standard.

At first glance one might conclude that this class is doing satisfactory work since the median score exceeds the standard established by the author of the test. But careful analysis of individual results shows tendencies to error, the correction of which would increase the efficiency of the class as well as that of particular individuals.

Nearly half the class found borrowing a source of difficulty. Thus, nineteen pupils apparently subtracted as if they had not borrowed. This inaccuracy gave rise to 31 errors. Seven pupils made errors due to subtracting as if borrowing had taken place when it had not. Even though some of these errors may be attributed to incorrect combinations, the analysis suggests that a number of this class might profit from some well directed practice drill in subtraction with borrowing.

On the surface this class tested in April appears weak in ability to subtract fractions. Example 24, $\begin{array}{r} 8\frac{7}{8} \\ 5\frac{3}{4} \end{array}$, was solved correctly but seven times; 25, $\begin{array}{r} 27 \\ 12\frac{5}{8} \\ \hline \end{array}$, not at all; 29, $\begin{array}{r} 75\frac{3}{4} \\ 52\frac{1}{4} \\ \hline \end{array}$, six times; 33, $^{5}/_{12}-^{2}/_{10}=$, once; 34, $\begin{array}{r} 6\frac{1}{8} \\ 2\frac{7}{8} \\ \hline \end{array}$, three times; and 35, $3\frac{7}{8}-1\frac{5}{8}=$, four times. However, thirty-four errors were made on examples 24, 29, 34 and 35, due *not* to inability to perform the necessary subtraction, *but to failure to reduce answers to lowest terms.* These were contributed by twenty-two pupils. Evidently reduction to the lowest terms had not been sufficiently emphasized by the teacher. The fact that no one solved example 25 seems to indicate that these pupils had not been taught how to subtract a fraction from a whole number. Seven pupils gave $23^{2}/_{0}$ as an answer to example 29. These children apparently do not know the significance of a denominator.

While, as Woody maintains, the tests are not safe criteria for the judgment of individual abilities, yet the performances of individual pupils are sometimes so markedly below or above the performances of the class as to furnish a reasonably accurate diagnosis of the achievement of these individuals. In this fifth grade there was one girl for whom drill on examples in subtraction of fractions such as represented in numbers 24, 29, 33, 34 and 35, would be largely waste of time since she solved all of them correctly. She solved thirty of the thirty-five examples correctly, but failed on three examples involving decimals, numbers 28, 30 and 31. Moreover, she was the one pupil in the class able to solve both numbers 26 and 27, examples in subtraction of denominate numbers. Her needs were not the same as those of the larger part of the class. She could possibly profit if permitted to take sixth grade arithmetic. Further evidence on the wisdom of such promotion in arithmetic could be found by consulting her record on the tests in addition, multiplication, and division.

A pupil may easily be misjudged by taking only his gross score into account. One boy had only six examples correct owing to the fact that he had multiplied in all examples possessing a single digit in the minuend.

Minor deficiencies of this class are indicated by errors made by three pupils who put down a cipher in subtracting a cipher from an integer, and errors by four pupils on example 10, $7-4=$.

3. A third grade class of 29 tested in addition.

The median for this class was 16.7 or 2.2 points above Woody's standard.

Among the first 16 examples the most troublesome were number 7, $3+1=$, missed 14 times; number 8, $2+5+1=$, missed 13 times, and number 14, $25+42=$, missed 21 times. A number of this class are clearly in need of reviewing this type of exercise. Eight pupils did not miss either of the three examples. These very likely need no drill on such processes. Ten gave "3" as the answer to example 7. Seemingly the + sign was mistaken for the sign of multiplication. "10" was given as the answer to example 8 four times, and "11" was given as the answer 3 times. The former may occur through multiplying the 2 by the 5, disregarding the 1, while in the latter case the process seems to be the same except that the one is added. "13" occurs three times as the answer to 14, and "31" occurs twice. The former is arrived at by adding all of the figures and the latter by adding the 2 to the 4 and then adding the result to the 25.

Only a very little attention on the part of the teacher should be necessary to remedy the weaknesses of the class in examples of this sort. The errors were due not to a lack of knowledge of the fundamental addition combinations, but to an incorrect mental "set." The bonds formed were very often correct for the process which the child set out to perform. As previously stated, distinction must be made between an error and the mental process causing the error. A number of the errors in this class are quite clearly errors in carrying. Pupil 12 fails entirely on this process. In six different examples each of which requires carrying he adds every column of each correctly with the exception that he disregards entirely the figures to be carried. He made no errors on examples requiring only simple column addition.

Pupil 24, a nine year old girl, differs radically from the other members of the class in the type of process she can manipulate. While she had only a median number correct, she was able to add correctly the fractions in examples 23, 25 and 28, which are $^1/_3+{}^1/_3=$, $^3/_8+{}^5/_8+{}^7/_8+{}^1/_8=$, $^3/_4+{}^1/_4=$, but did not reduce to lowest terms in the latter two. She solved also example

24, $\begin{array}{r}4.0125\\1.5907\\4.10\\8.673\\\hline\end{array}$ involving decimals and example 20, $\begin{array}{r}\$12.50\\16.75\\15.75\\\hline\end{array}$ in U. S. money. Her errors on two other examples in addition of U. S. money, (numbers 19 and 21), were due to incorrect addition and not to difficulties offered by U. S. money. She is a pupil whose score on a test of this sort would be raised by an acquisition of the ability to recognize the sign of operation in examples 7, 8 and 14. Seven pupils in this class solved example 24 in addition of decimals, and nine solved example 21, $\begin{array}{r}24\\2\frac{1}{2}\\\hline\end{array}$. Future teachers of this class should ascertain which of the pupils will need only a very limited additional teaching to make them proficient in addition of examples in decimals and in U. S. money.

4. A fourth grade of 30 pupils tested in addition.

This class scored 21.7 or 3.4 above the Woody standard.

Examples 7, 8 and 14 again offer a difficulty to this class in recognizing the sign of operation. Seven errors were made on these three examples. Three pupils made two each. These children show signs of weakness in examples of this type. The record sheet of individual scores by examples reveals some pupils already strong in addition of fractions and decimals, and for whom indications seem to point that it would be a waste of effort to spend time later in developing these processes. Examples like number 23 or number 25, are not too difficult for four of this class. Sixteen, or more than half of the class, solved number 22, an example in addition of three long columns of figures. Pupil 3 added correctly every example in decimals including those in U. S. money. Pupil 13 on the other hand, while able to solve the three examples in U. S. money, failed on examples 24, 31 and 33 in decimals. But this same pupil had examples 23, 25 and 28 in addition of fractions correct. He appears to be able to add fractions of the same denomination. The highest score in the class, 27 correct, was made by pupil 14, a nine year old boy who has already skipped a half grade. He had examples 23, 25, 28, 29 and 32 in addition of fractions correct. In fact, he made only one mistake on the first 25 examples. He was charged with an error on example 2 on account of omitting the dollar sign. Little further teaching effort will be needed to make him proficient in addition of fractions and decimals.

The poorest score in the class was 13 correct made by a 14 year old girl. She missed examples 8, 11, 14, 16 and 17, solved example 18 correctly, and had none correct from there on. Her answer to example 8, $2+5+1=$ was "11." Her answer to example 14, $25+42=$ was "49," obtained by adding the 2's and then adding the 4 to the 5. In the other examples attempted and missed, up to number 23, the error was in addition. Her difficulty appears to be in carrying, for the first column in 5 of the 7 incorrect examples in column addition was added correctly.

The illustrations of class diagnosis so far given are only typical of what may be discovered through the Woody tests. Additional illustrations might be given for each grade in each of the four fundamentals. Those already given should be sufficient to indicate the type of service which these tests may perform. Through the use of the tests and by an analysis of the performances on them, weaknesses in particular types of examples may be discovered. Not infrequently it is possible to discover the particular mental process operating in the minds of children to prevent the correct solution of a certain type of example. Often unusual proficiency is revealed and methods of further development can be discovered. The need of differentiated drill which shall provide one type of practice exercise for some of the class and another for other members, even though all may be about equal, can often be brought out. Such discoveries furnish points of attack for the teacher and the supervisor and make for increased efficiency in teaching. The illustrations in class diagnosis given above show that even though a class may be above the standard for its grade, very often marked weaknesses in simple processes are present, which if remedied would permit of a much higher score.

From the standpoint of a superintendent or a supervisor it is often desirable to locate weaknesses which may be rather common for the system as a whole. For this purpose a test must be sufficiently brief so that an undue amount of clerical labor shall not be required to discover the desired facts. At the same time the test must be sufficiently comprehensive to include a variety of types of examples in order that a reasonably detailed diagnosis may be made. Both of these requirements are met by the Woody tests. How the tests may be used by a supervisor to discover something as to the teaching needs in a number of schools of the same system may be judged from the two tables following. Table V, from the survey of the

TABLE V

Ten Troublesome Examples in Grade VIII

Examples	Standard*	Percent of pupils who solved each of ten examples correctly—for the city and by buildings					
		Buildings					
		City	1	2	3	4	5
$2 \div 2 =$	86.7	71.9	64.5	61.1	89.7	71.4	69.7
$2.1\overline{)25.2}$	79.9	67.6	61.3	66.7	65.5	75.	69.7
$25\overline{)9750}$	87.8	77.7	74.2	77.8	79.3	82.1	75.8
$.003\overline{).0936}$	73.2	53.2	35.5	55.6	65.5	46.4	72.7
$3\frac{1}{2} \div 9 =$	59.	33.8	35.5	5.6	34.4	42.9	39.3
$\frac{3}{4} \div 5 =$	57.7	51.1	48.4	27.8	51.7	64.3	54.5
$^5/_4 \div {}^3/_5 =$	59.6	33.8	38.7	16.7	34.5	32.1	39.4
$52\overline{)3756}$	46.9	45.3	38.7	61.1	41.4	57.1	36.3
$531\overline{)37722}$	25.1	28.1	25.8	33.3	27.6	32.1	24.2
$9\overline{)69 \text{ lbs. } 9 \text{ oz.}}$	25.7	18	35.5	5.6	3.5	25.	15.
No. pupils taking test		139	31	18	29	28	33

Janesville, Wisconsin, schools,† gives the per cent. of eighth grade pupils in each building who solved each of ten examples of the division scale correctly.

Possibly few eighth grade teachers would be willing to admit previous to a test of this sort that their pupils could attain averages on these examples no greater than did the 139 children shown in Table V. Through the application of the test a supervisor can avoid personal criticism of a teacher for her poor success. It becomes unnecessary as the facts speak for themselves. There are perhaps few teachers who would not be stimulated to immediate efforts to remedy the shortcomings of their pupils here revealed. No true teacher would knowingly ignore a condition where only 72% of the pupils in the eighth grade solve examples like $2 \div 2 =$. Yet this was an undiscovered condition existing in this city previous to the test. Many gave 2 or 0 as answers. That an example like $^5/_4 \div {}^3/_5 =$ was beyond two-thirds of the eighth grade or that

*Computed from Table XXX of Woody's monograph. See Tables III and XXXI for key to corresponding orders of examples.

†*An Educational Survey of Janesville, Wisconsin.* State Dept. of Education, Madson, Wis.

only 45% of them could solve the example $52\overline{)3756}$ correctly was perhaps undreamed of by teachers and supervisors in this city. None of the buildings of Table V performed as well as one might desire on any of the examples, even though certain of them exceeded Woody's standard in some instances.

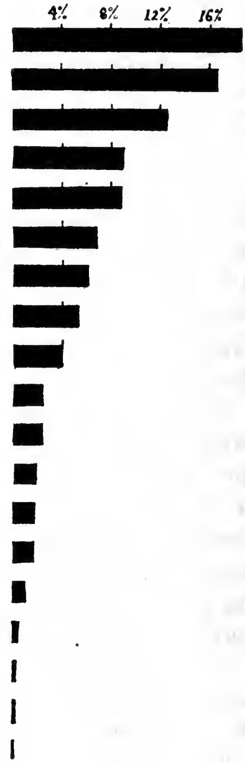
Table VI (Figure III), also taken from the Janesville survey, is a classification of 1500 errors on examples attempted in grades seven and eight on the division scale. These errors were contributed by 181 pupils, 43 from grade seven and 138 from grade eight. All of the eighth grade papers were first included and sufficient seventh grade records were then added to make the necessary 1500. This resulted in the selection of one seventh grade of 30 pupils and the first 13 pupils whose papers were drawn from a second.

No account is taken here of the difference in the number of possibilities for making each type of error on the division test nor of the omissions. A record of these facts would add to the precision of the diagnosis. There are many chances for making errors in simple subtraction, multiplication, division and placing, but comparatively few for making certain others. Opportunity for applying the process of inverting the divisor is offered in only five examples and decimals occur in five, though they may be used in some others. Decimals and inversion of fractions, however, cause a large proportion of failures. Evidently the prescription of drill work for these children should center more on these processes than on others.

Examples 29, 30, 31, 32, and 34 of the division scale offer opportunity to apply the process of inverting. These examples are: $3\frac{1}{2} \div 9 =$, $\frac{3}{4} \div 5 =$, $\frac{5}{4} \div \frac{3}{5} =$, $9\frac{5}{8} \div 3\frac{3}{4} =$, $62.50 \div 1\frac{1}{4} =$. 277 errors out of 905 (181×5) opportunities were made in this process. In addition 68 omissions were recorded. There were 187 cases of incorrect solutions on account of other errors in the work or on account of errors in other operations which pupils attempted to substitute for inverting—the most logical process to use—*e.g.*, in example 34 several reduced the divisor to a decimal, 1.25, and then erred in division of decimals. Only 7 of the 181 children made no errors of any kind on either of the five examples. If we add the 277 and 68 (which we may do since ample time is allowed for the test), we have at least 345 errors that may fairly be charged to insufficient facility with the process of inverting. That is 38 per cent. of the opportunities were failures on account of omission or inability to manipulate the process. If

Table VI. FIGURE III

	No.	%
1. Incorrect Inverting or Failure to Invert	277	18.46
2. Incorrect Division	246	16.4
3. Incorrect placing, of, or omitting of decimal point	189	12.6
4. Failure to reduce or incorrect reduction of remainder or answer	133	8.87
5. Incorrect division of denominate numbers	132	8.8
6. Omitting or adding cipher in quotient	101	6.73
7. Not completing division	92	6.13
8. Incorrect subtraction	81	5.4
9. Incorrect multiplication	61	4.07
10. Using incorrect denominator of remainder	36	2.4
11. Failure to classify result	35	2.33
12. Failure to employ both parts of fraction	29	1.93
13. Incorrect placing or bringing down	28	1.87
14. Incorrect reduction of mixed numbers	28	1.87
15. Failure to recognize the sign of operation	15	1.
16. Incorrect addition	7	.47
17. Error in copying	4	.27
18. Illegible answer	4	.27
19. Incorrect cancellation	2	.13
Total	1500	100%



time permitted one might readily determine how many of the 187 other errors occurred in substituted processes. It is safe to say that at least 40 per cent. of the total chances on these examples resulted in failure either directly or indirectly because of insufficient familiarity with the process of inverting. These are significant facts for the supervisor. The diagnosis has revealed one marked weakness in arithmetic work in this school system. If desired, one might record the percentage of error separately by classes, *e.g.*, in the seventh grade class of 30 pupils there were 58 errors in attempted use of the process of inverting, and 14 omissions, making an efficiency of but 52%, $(150-72) \div 150$, in the attempted use of this process. The errors were contributed by 23 pupils. Three pupils made 5 errors, four made 4, three made 3, five made 2, and eight made 1. In one eighth grade of 31 pupils there were 42 failures contributed

by 23 pupils in attempted use of the process of inverting, and 16 omissions or an efficiency of 62%, (155 − 58) ÷ 155. While the showing is still unsatisfactory, it indicates a higher proficiency than that of the seventh grade above.

Omission of a cipher in the quotient was another deficiency of the group of children who contributed the errors of Table VI. Examples 18, 21, 23, 24, and 25, 13)65065, 25)9750, 23)469, 75)2250300, 2400)504000, involve ciphers in the quotient. The 181 children made 101 errors of this type. In the seventh grade group of 30 pupils 9 made no errors by reason of omitting a cipher in answer to one or the other of these examples, another added an extra cipher in one, and four others omitted one or more of the examples entirely. Ten of the grade appear to be definitely weak in examples involving ciphers in the quotient. These ten failed on two or more of the five examples either through omitting a cipher or omitting the example altogether. The performances of each of the other classes on these examples might be analyzed similarly.

Example 23 serves as a good index not only of whether pupils have mastered the difficulties offered by a cipher occurring in the quotient, but of whether they have formed the important habit of estimating answers. Six of the thirty pupils in the seventh grade class above gave $2\,{}^{9}/_{23}$ as an answer. Evidently they have not been taught to estimate answers habitually. As we have stated before, it is not so much a question of how many errors, but what do the errors reveal. "It isn't the fact that you're dead that counts, but how did you die."

That a single example may sometimes be sufficient to reveal distinct sources of weakness, may be seen from the results for this same seventh grade class on example 28 of the division test, .003).0936. The errors on this example reveal an unsatisfactory development in ability to divide one decimal by another. Only nine of the class had the correct answer 31.2. Seventeen failed because of inability to manipulate the necessary operations. Three others omitted it entirely apparently for the same reason. One other made what seems to have been a careless error in expressing the quotient as 21.2. Only two of these 17 errors could possible be explained on the basis of being due to a careless omission of the decimal in the quotient. These obtained 312 as an answer. Two gave an answer .0000312 (obtained by pointing off the sum of the places

in the dividend and the divisor), one .00312, five .0312, one 301.2, one .12½, one .123, one 13, one 2, one .209 $^{9}/_{003}$, and one .9. The last six of these were obtained through an attempt to solve the problem in some such manner as this:

$$
\begin{array}{r}
.12\tfrac{1}{3} \\
.003\overline{).0936} \\
\underline{.003} \\
906 \\
\underline{.006} \\
900 \\
300 = \tfrac{1}{3}
\end{array}
$$

It is significant that 16 of the 17 above also missed example 20 $2.1\overline{)25.2}$ but did not fail on either examples 22 or 26 which are: $2\overline{)13.50}$ and $12\overline{)2.76}$. Eight of the nine who had example 28 correct had also 20, 22, and 26 correct. Of the three who omitted example 28 one omitted all four of the group and the other two solved examples 22 and 26 correctly but failed on number 20. The boy who erred in division on example 28 solved the other three correctly. Thus at least 19 of this class appear to be weak in dividing one decimal by another of the type in examples 20 and 28 and 9 appear strong.

In his criticism of the Woody scales the writer of the article referred to in Part I omits all analysis of the types we have just shown. Yet it will be recalled that he stated as his purpose, "to learn how completely it is possible for these scales to diagnose a class and how accurate is such diagnosis as they make." He attempted to prove his major thesis that the tests do not afford a satisfactory diagnosis by offering certain objections. His principal objections were that the scales do not contain all types of examples or a sufficient number of those types; and that the statistical rather than the analytical method was used in their selection. His supporting evidence was limited almost entirely to a comparison of the percentage of examples correct for given types with figures showing the average number of similar examples solved on the Cleveland Survey Tests,—a rather faulty comparison at best. We submit the illustrations which we have just given in Part II of this paper as evidence in favor of the Woody Scales which can not be overlooked in any consideration of their diagnostic merits.

DISCUSSIONS AND COMMUNICATIONS

SOME RESULTS WITH THE HAHN-LACKEY SCALE IN GEOGRAPHY

During the past few months, the writer, with the assistance of two of his graduate students who are engaged in teaching geography in elementary schools, has been able to make some studies with the Hahn-Lackey scale.* No opportunity has been found as yet to extend the study over many schools, but it is expected that this may be done eventually.

The making of a scale, and, no less, the use of one, in geography, offers problems of peculiar difficulty. It would seem that only through careful and repeated testing of the scale, can any real progress or betterment in its construction be brought about.

The objects of the work done with the scale were:

(a) To stimulate geography teaching by emphasizing some minimum essentials.

(b) To enable certain students of teaching to form some general notion about the quality of geography teaching in their schools.

(c) To study the scale itself by using it.

Schools Giving the Test

For convenience, the present report will give the results noted in three schools, to be designated hereafter as A, B, and C, respectively.

School A is a good sized school in an eastern city. Thirty-seven per cent. of its pupils are foreign born—Italians and Russian Jews. 460 children in grades four to eight were tested, using steps M, P, Q, S, W, U and V of the scale.

School B is a very large school of 3600 children located in the same city as *A*. More than one-third of the children are Hebrews. One-tenth are of Italian birth. In this school there are from seven to ten classes in each grade, therefore there was ample opportunity to test children of supposedly equal attainments. 1595 children, in grades four to eight, were tested with the very simple steps W and X of the scale. 799 pupils were tested with step S.

School C is a summer school for children in a college town in the East. These children were of typical American stock. 116 pupils in grades four to seven were tested.

The writer would direct attention to the fact that the present report is based on answers written by 2970 children. In making the entire scale, papers from 1696 children were used by Hahn and Lackey.

*Vide Jour. Geog., Vol. XIII, No. 5, January 1918.

STEPS OF THE SCALE USED IN THIS INVESTIGATION

Step M

90. How could you go to Asia if you wished to make that trip?
93. Name two large rivers of Asia.
111. How can one go by boat from the Hudson River to Lake Erie?
133. Give two reasons for the importance of the Columbia River.
136. Name two of the most important materials shipped on the Great Lakes.
155. Give the capitals of Japan and China.
159. Name five important inland cities of Europe.
212. Draw a map of your state, and locate in it two rivers, the capital, and the largest city.
214. Name the state or territory in which each of the following is located: Galveston, Washington, St. Paul, Sitka, Savannah, Spokane.
199. What disadvantage do the people of Great Britain suffer as to food supply?
109. Why do the rivers of New England furnish water power for manufacturing?
135. Why is the Rio Grande an important river?
138. Give two reasons why cities usually grow up at falls in rivers.
146. Give two reasons why Argentina exports wheat to Brazil rather than to the United States.
165. What part of Asia is similar to Canada and in what way is it similar?
188. Since the larger part of our iron ore is mined in Minnesota, why is little iron and steel manufactured there?
117. Why does the earth not look round to us?
76. Why are so many hogs raised in the United States?

Step P

79. Name two large rivers in South America.
80. What two important products are brought to us from Brazil?
88. How can you get from New York to London, and in what direction would you go?
89. How can steamboats go from New York City to San Francisco by the shortest route?
100. Name the chief occupation of the people of Australia.
105. Give two reasons why mountainous regions are not good for farming.
28. How can we tell how big a country is by studying a map?
10. How long does it take the earth to go around the sun?
67. Give the principal reason why such dense forests grow along the Amazon River.
70. Which part of the United States is most important for manufacturing and which for agriculture?
74. Where may snow be found in the hot belt near the equator?
77. Why is there so little rainfall in the Great Basin of the U. S.?
78. Name one way in which the Panama Canal will be an advantage to the United States in its trade with South America.
81. During what months does Argentina have winter?
82. In what industry or kind of work are most of the people of England employed?
31. Why do so few people live in deserts?

Step Q

60. Name four fur-bearing animals.
36. Name four large rivers of Europe.
99. What two countries of Asia are noted for tea?
129. Give the capitals of Colorado and Massachusetts.
42. Name two kinds of food that we get from animals.
1. From what country do we get much of our coffee?
12. Which is the largest and which is the smallest, the moon, the sun, or the earth?
71. What country of South America has a climate simliar to ours?
103. Why is it warmer in New Orleans than in Chicago?
104. Why are camels such useful animals for traveling in desert places?
1. What is the cause of day and night?
18. What do the rivers do with the soil that they carry?

Step S

52. What is the largest city of your state?
64. Where is Alaska and to whom does it belong?
84. Name four large cities of Europe.
92. Give the capitals of France and Germany.
101. Name two large bodies of water that border on Florida.
45. Name four things you use for food that do not grow where you live.
68. Give one reason why so many of the great cities of the United States are near the sea coast.
72. Which is the coldest and which the warmest part of South America.

Step U

63. What country is north of the United States and to whom does it belong?
102. Name two other countries in North America besides the U. S. A.
24. Name five wild animals.
5. What two oceans border on the United States?
43. Name a plant used for making cloth.
49. Write your whole aRdress.
27. Name two kinds of work that men do in getting materials for building houses.
26. Name two kinds of work that men do in getting food for us.
34. How can you tell from what direction the wind is blowing.
50. To whom do the streets or roads belong?

Step V

59. In what direction do you live from the equator?
64. Name two important mountain ranges of the United States.
11. Name the four reasons.
33. What is the direction half way between south and west?
8. In what direction would you go to go to Canada?
15. What is the capital of the United States?
25. Name an animal useful to man in desert countries.
65. Why is there not much farming in Alaska?
22. How do we know that there is air?
4. Why is the Arctic Ocean not much used by sailors?
40. How does the ocean help to furnish us food?
30. Why are there more birds here in summer than in winter?

Step W

32. In what direction are you facing when your back is toward the north?
51. What is the capital of your state?
17. Name two things that plants must have to live.

Step X

2. Name two animals used by the Eskimos.

NOTE: All numbers prefixed to the above questions correspond to the numbers given to the questions in the original scale as published by the authors.

RESULTS OF THE TESTS

Below are given the tabulated results as calculated for the three schools.

SCHOOL A

Grade	Step	Number of Pupils Taking Test	Per cent. Obtained	Per cent. called for on scale	Remarks
8th	S	33	86.1	88	General failure on question 45.*
8th	S	29	72 3	88	Did poor;y on question 68.*
8th	Q	32	78.2	79	Failed on rivers of Europe.
8th	P	26	65.5	73	Nearly all failed on question 81; many on 105.*
8th	M	34	68.4	50	
8th	M	33	57.8	50	
7th	P	30	60.4	73	Nearly all failed on questions 81;* many on 105.*
7th	P	28	57.1	73	Ditto
7th	M	32	43.7	50	
6th	P	40	48.8	58	Nearly all failed on question 81;* many on 105.*
6th	M	38	34.7	34	
5th	U	30	74.0	84	Questions 26* and 27* were general iailures.
4th	W	38	82.4	84	Failed on food plants need
4th	V	37	76.2	79	Failed on 65* and 22.*

Total number tested—460 pupils.

*For these questions see above.

SCHOOL B

NOTE: A and B under "Grade" mean respectively the first half and second half of the year indicated. Results with step X (1595 papers rated).

Grade	No. getting zero	No. getting 50%	No. getting 100%	Total No. of Pupils	Average percent obtained	No of Papers	Average % obtained	Percent called for on scale
8B	0	14	61	75	90.6	210	87.1	99
8A	6	52	87	135	83.6			
7B	2	35	115	152	87.0*	268	82.7	99
7A	6	38	72	116	78.4			
6B	26	59	90	175	68.2	321	66.9	96
6A	17	66	63	146	65.7			
5B	46	101	72	219	59.9	382	55 2	94
5A	28	92	43	163	54.6			
4B	43	96	79	218	58.2*	414	53.8	88
4A	44	110	42	196	49.4			

*Step up, not down.

No deduction made for incorrect spelling.

Did not accept "deer," "wolf," "bear," "polar bear" or "whale," although these were commonly given by children. (Perhaps part credit should have been given for "deer" especially, and "wolf" meaning "dog.")

RESULTS WITH STEP W. (1595 PAPERS)

Grade	*Number of Pupils Getting* 0	16⅔%	33⅓%	50%	66⅔%	83⅓%	100%	Total No. of Pupils	Aver. % obtained	No. papers	Aver. % obtained	Percent. called for on scale
8B	0	0	0	1	5	23	46	75	92.0	210	93.2	98
8A	0	0	0	8	14	53	70	135	94.4*			
7B	0	0	1	5	10	51	85	152	90.1	268	89.7	98
7A	0	1	0	7	8	31	69	116	89.3			
6B	0	2	3	10	26	47	87	175	85.6	321	83.8	94
6A	0	3	3	17	8	62	53	146	82.1			
5B	1	9	5	36	29	68	71	219	76.7	382	74.1	92
5A	0	9	12	31	29	34	48	163	71.5			
4B	0	10	5	69	72	43	19	218	64.5	414	54.1	84
4A	12	30	42	50	61	0	1	196	43.7			

*Step up, not down.
No deduction made for incorrect spelling.

RESULTS WITH STEP S. (799 PAPERS)

8B, 8A								210	90.8	88
7B, 7A								268	89.7	88
6B, 6A								321	84.4	79

SCHOOL C (116 PAPERS)

Grade	Scale step	*Number of Pupils Getting* 0	16⅓%	33⅓%	50%	66⅔%	83⅓%	100%	Total No. of pupils	Average % obtained	Percent called for on scale
7	W	0	1	0	0	0	1	1	3	66.6	98
6	W	0	0	1	5	5	4	3	18	69.4	94
5	W	0	1	0	5	7	1	4	18	67.5	92
4	W	1	1	2	9	5	0	1	19	50.8	84
7	X	0			2			1	3	66.6	99
6	X	0			7			11	18	80.5	96
5	X	1			11			6	18	63.8	94
4	X	2			11			6	19	60.5	88

Comments and Conclusions

As usual in this kind of work there were many strange answers. Some examples: Question—Name two animals used by the Eskimos. "Wales and popular bears;" "dogs and flanders;" "dogs and antler;" "iron and ore." Question—Name the capital of your state. "Hudson River;" "the government;" "Lincoln;" "President Wilson;" "San Francisco;" "United States;" "Alabama;" "the first letter of your name;" "meat." Question—Name two things that plants must have in order to live. "Dairying and manufacturing;" "potatoes and onions;" "coal and light;" "cabbages and potatoes;" "poison ivy and summer salt."

We cannot be too careful in recasting and polishing our questions so as to bring them within the range of every child who really has the knowledge to answer properly if the questions are clear.

The data in hand do not warrant any sweeping conclusions.

I believe, however, that the scale does offer a more accurate method of testing, and of measuring results, than the usual geography examination. At the same time our experience would seem to indicate the advisability of altering the phrasing of some of the questions along at least two lines: (a) to bring forth more specific answers, (b) to insure the children's grasping the meaning. As an example of "a" I would cite the scale question: "Name two ways in which the farmer helps to get food for us." As an example of "b" I would cite the scale question: "Name two things that plants must have to live."

The language difficulty has not been eliminated from this scale.

The results from the use of the scale disclosed marked deficiencies on the physical geography side. The Boston Geography Report brought out similar deficiencies in the children's knowledge. It would appear that possibly too much has been expected or attempted on this line with children.

The table for Step W, School B, shows a surprising regularity of increasing geographic content in the minds of the 1595 children tested. Even though they stood below the scale standard at all stages, nevertheless their progress was quite satisfactory.

On the whole, there is apparently a bit of truth in the statement that the scale is as much a test of general information as of geography. The fact that, as a rule, the results with eighth grade children were relatively nearer to the scale per cents. than were the results in lower grades gives support to this idea.

Specific weaknesses uncovered in the schools tested were a knowledge of the native state, and of certain features of Europe, notably the rivers. Teachers have in this connection spoken of the great practical value of the scale in showing them where to lay more emphasis in their work.

Further studies are being made and it is hoped that new light on the value of the scale will be obtained.

CHESTER A. MATHEWSON.

Brooklyn Training School for Teachers.

The Journal of Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Published monthly except July and August by
WARWICK & YORK, INC.
HERSHEY, PA. BALTIMORE, MD.

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y.

NOTICE TO SUBSCRIBERS. Notify the publishers, Warwick & York, Inc., of any change of address; as copies lost in transit because of subscriber's removal to another address will not be duplicated.

THE JOURNAL is issuing a series of Monograph Supplements for the publicaiion of more extensive researches in educational psychology. For details prospective contributors may apply to H. E. Buchholz, 10 East Center St., Baltimore, Md.

The price of the JOURNAL is $3.00 per year of ten monthly issues, payable in advance. Single issues are 50 cents each. Back volumes are $5.00 in original wrappers and $6.25 bound in library buckram. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME IX DECEMBER, 1918 NUMBER 10

EDITORIAL

The Gary Survey

Few school systems have received such wide advertising as that of Gary, Indiana. Few school surveys have aroused such widespread interest and such keen anticipation among all classes of educators in this country as the survey of that system undertaken by the General Education Board. And rarely indeed does a school survey afford such an opportunity as is now offered the many practical school men who have visited Gary to say "I told you so." What is it that we find? In the first place it is obvious that we have here the results of one of the most careful, painstaking and thorough surveys that has ever been made. The surveyors were recognized authorities in their fields and most of them had had extended experience in studying school systems. No expense was spared, and time enough was taken in the examination of the system for the surveyors to become thoroughly acquainted with the details of its workings. The most approved educational tests were used in the study of classroom products, and the greatest pains were taken to obtain objective and unbiased judgments.

And what are the final pronouncements of the survey on the efficiency of the organization as judged by the results attained? On the one hand the people of Gary are given great credit for adopting a broad conception of education and providing for a greatly enriched curriculum. The attempt has been made in the Froebel and Emerson schools to supply a physical equipment adequate to the demands of progressive educational ideals. This, of course costs money,—much more money than the traditional school building equipped with its stereotyped furnishings for a few abstract bookish subjects. Again it is found that the Gary plan of organization can accommodate a larger number of pupils than the same plant operated on the conventional plan. This is something of a truism. The real question is whether the Gary plan produces as good results in the attainments of the pupils as the conventional plan. Finally Gary has attempted to practice democratic theory in school conduct and discipline. This results in an atmosphere of freedom from restraint, an independence, a self-reliance sometimes bordering on self-assertiveness, that is not favored in the conventional school. What of its ultimate effect on the social and moral development of the pupils?

On the other side of the ledger the survey is perfectly frank in pointing out the weaknesses and defects of the Gary system,—and these are so serious as to give the thoughtful educator occasion to hesitate. Much has been made of the auditorium, the aesthetic and entertaining features of the Gary schools, and these have been undoubtedly well conducted. But there are some things in any school that have to be learned. One of the chief values of educational procedure is the demand it makes of the pupil to concentrate all his intellectual resources and *master* the difficulties that he encounters. Here is where the Gary atmosphere fails. The shops and the classrooms, as well as the auditorium and the playground, seem to be places for amusement and entertainment, the teachers work hard to supply diverting activities, and the pupils undertake what they feel like doing and enjoy themselves as much as they may. The administrative control of the Gary schools is so weak as to be almost non-existent. No school system can be trusted to run itself, much less a system embodying so many novelties as Gary professes. The educational supervision is defective. The standards of performance whether in the workshops or in the classrooms are distinctly low. Not that the principals and teachers do not want good work done, but the painstaking attention to details by which alone good work can be obtained is too irregular to be effective. Slipshod records are allowed to go unchallenged, self-criticism is lacking, and projects which seem to be ideal in theory fail for lack of careful scrutiny of results.

The study of the results produced by the Gary plan is therefore distinctly unfavorable. The scores obtained in the classroom tests showed the usual amount of variability from class to class, and while

some rooms made high records the average results fall short of legitimate expectations. The report is inclined to defend the theoretical aspects of the plan and to account for the poor showing in results on the basis of careless administration, lack of educational supervision, and a resulting lack of thoroughness and intensity of effort in the classroom work. But the world has long been familiar with the theory and has looked to Gary for the demonstration of its efficacy in practice. Now, under the most friendly examination, satisfactory results are lacking. May it not be that the very fundamental *laissez-faire* principle of the Gary idea is at fault? Are there not certain obligations resting upon teachers and pupils alike in their relations to each other and to school work that are essential to securing educational results? Is it not of the very essence of education to build up such habits of surmounting difficulties in arithmetic and history and geography and spelling that pupils will stand to their guns and in spite of distaste and discomfort will push through to the goal of attainment? Perhaps it is the very flabbiness of moral fibre inherent in the Gary theory that has caused the administration to be so weak, the supervision so inadequate, and the teaching so spineless.

J. C. B.

NOTES AND NEWS

The November meeting of the New York Society for the Experimental Study of Education had as its general theme the subject of "School Surveys." Principal Oswald Schlockow, of Public School 50, Brooklyn, reported on "A Study of the Pupis of Public School 50," which was undertaken at his request by the Division of Reference and Research shortly after he came to the school last spring. In the report stress was laid upon the results obtained from the use of the Woody Tests in Arithmetic, the Trabue Completion Tests, and the Briggs Cycle Test in English. Attention was called to the value of such surveys for teachers and principals, and the desirability of having a permanent organization of trained investigators in a city school system to carry on such surveys. Superintendent Don C. Bliss, of Montclair, New Jersey, gave an account of some testing that is being carried on in the Montclair schools, and Dr. Benjamin R. Simpson, of the Brooklyn Training School for Teachers, spoke on "Some Aspects of the Janesville, Wisconsin, Survey." At this meeting the Society organized by adopting a constitution and electing the following officers: President, William E. Grady, District Superintendent; vice-president, Stephen F. Bayne, Principal Public School 5, Manhattan; secretary-treasurer, J. Carleton Bell, Brooklyn Training School for Teachers; additional members of the executive committee, Arthur C. Perry, Jr., District Superintendent, and Truman Lee Kelley, Teachers College.

At the annual meeting of the American Psychological Association, held at Baltimore, December 27 and 28, the papers were limited to psychological work in connection with the war. At a joint session with Section H (Psychology and Anthropology), of the American Association for the Advancement of Science, a symposium was presented upon "The Future of Pure and Applied Psychology." Professor E. L. Thorndike, retiring vice-president of Section H, delivered an address on "Scientific Personnel Work in the United States Army," and Professor E. F. Buchner, retiring vice-president of Section L (Education), took as the topic of his address "Scientific Contributions of the Educational Survey."

The annual conference of the Education Association of Western Pennsylvania and of the Pittsburgh City Institute was held in Pittsburgh, November 29 and 30. The main topic was "The Results of Measurement," and one of the principal speakers was Mr. Stuart A. Courtis, director of the Bureau of Research, Detroit City Schools.

The first issue of the *Social Hygiene Monthly*, published by the War Department Commission on Training Camp Activities and the United States Public Health Service, has recently come to hand.

The journal is designed to assist the Government in its nation-wide crusade against venereal diseases. Copies may be procured by writing to 105 West 40th Street, New York, N. Y.

St. Louis is planning to establish a laboratory for the mental testing of children. Part of the money for the maintenance of the laboratory is to be furnished by the council of social agencies and part by the public schools. The plan is in every way like that governing the psychological laboratory in Cincinnati.

Edward Kidder Graham, president of the University of North Carolina, died of pneumonia at his home in Chapel Hill, N. C., on October 26. President Graham has stood for the democratic organization of university administration, and for the utmost possible service of the university to the people of the state. He was a pioneer in the work of university extension in the South, and through it became the guide and leader of all the forces within the state working for the fuller and freer life of the people.

President Charles R. Van Hise, of the University of Wisconsin, died on November 18. President Van Hise was one of the leaders of educational thought in the Middle West, and was a strong advocate of a national research university to be located at Washington.

Professor Morton Prince, founder and editor of the Journal of Abnormal Psychology, has been elected corresponding member of the Societe des Hopitaux de Paris.

Lieutenant George O. Ferguson, Jr., associate professor of psychology of Colgate University, is stationed at Camp Lee, Virginia, in care of the psychological examination of men at that camp.

Dr. William T. Sanger has been made dean and professor of psychology and education at the State Normal School, Harrisonburg, Virginia. He has served as director of the summer session for the last two years.

At the University of Washington Associate Professors Leonard V. Koos and Fred C. Ayer of the college of education were recently promoted to full professorships.

Dr. Allen J. Thomas, instructor in philosophy in Cornell University, has been appointed professor of philosophy, psychology and education in the Connecticut College for Women.

Dr. Wilford E. Talbert, director of the Bureau of Educational Investigation, San Francisco, has been appointed research assistant to Professor G. M. Whipple, who has charge of the Bureau of Salesmanship Research at the Carnegie Institute of Technology, Pittsburgh.

Among the books announced for early publications by Longmans, Green and Company are *Education and Social Movements, 1700–1850*, by A. E. Dobbs, and *An Introduction to General Physiology*, by William Maddock Bayliss.

PUBLICATIONS RECEIVED

DON C. BLISS. *Methods and Standards for Local School Surveys.* New York: D. C. Heath and Company, 1918. Pp. xxiv, 264.

The school survey has been the focal point during the past eight years for the application of educational measurements. Whether this has been an unmixed blessing for the science of education may be open to question, but it has undoubtedly greatly stimulated the spread of the measurement ideal. The results of these surveys have not always been readily accessible, and the author of the present book has rendered a significant service in bringing together the most important findings of these investigations. As Professor Strayer says in his introduction, "Superintendent Bliss has given us in this book a manual which will be of very great help to the superintendent of schools who wishes to attack his problems in a scientific manner. He has brought together from his own practice and from the field of school surveys and investigations in educational administration the methods which can be most satisfactorily employed in checking up the work of a school system. He has dealt with the problems of attendance, classification and progress of children, with the financing of the school system, with the necessity for a building program, with the technique of measuring the achievements of pupils, and many other questions which are continually before the superintendent for solution. Throughout the book he has presented the case just as the superintendent of schools has to meet it from day to day. The method of attacking each of the problems has been described with such directness and simplicity as to make it available even for those who have had little special training in administration."

LEWIS CARROLL. *Alice's Adventures in Wonderland.* Edited by Clifton Johnson. Cincinnati: American Book Company, 1918. Pp. 154.

A richly illustrated edition of this children's classic designed for use as a supplementary reader in the schools. It should have a wide circulation.

CHARLES S. GARDNER. *Psychology and Preaching.* New York: The Macmillan Company, 1918. Pp. 389. $2.00.

"The field of educational psychology has been very thoroughly worked over, though the last word has by no means been said. The help which teachers have derived from it is very great, and no one now is considered equipped for that noble profession who has not mastered its principles. But so far as my knowledge extends. there have been few serious efforts to apply modern psychology to preaching . . . But there seems to be just as much reason for applying the principles of modern psychology to preaching as for applying them to teaching." Nor is it a mere collection of cursory reflections on the experiences of preaching that the author offers us. On the contrary it is a serious, thoughtful, scholarly work whose every page reveals the breadth of his reading and the soundness of his scientific foundations. Not only is the book a distinct contribution to the ever widening field of applied psychology, but the author's frank acceptance of the attitude of scientific inquiry even in matters of religious belief makes the work a valuable addition to the literature of religious sociology.

KENNETH SCOTT LATOURETTE. *The Development of Japan.* New York: The Macmillan Company, 1918. Pp. xi, 237. $1.50.

This admirable interpretation of the history and present conditions of the "Island Empire" will be welcomed with gratitude by all those who are interested in things eastern. The narrative is direct and compact, but clear and readable. The account of the external and internal development of the country in the past ten years affords a good foundation for appreciating Japan's present position in world politics.

E. MARKHAM LEE. *On Listening to Music.* New York: E. P. Dutton and Company, 1918. Pp. viii, 159. $1.50.

Many intelligent people who have a good native basis for musical appreciation fail to get much out of music because they do not know how to listen. The present work affords a sort of general introduction to musical analysis. After two general chapters on how music is constructed, the author takes up the chief forms of music, such as the orchestral compositions, the instrumental concerto, chamber music, pianoforte recital, the vocal recital, the oratorio, madrigals and part-songs, church music, organ recitals, the opera, and miscellaneous forms of concert, sketches the history and aim of each form and gives typical illustrations. For readers of some musical training the book should prove very interesting and useful.

JACQUES LOEB. *Forced Movements, Tropisms and Animal Conduct.* Philadelphia: J. B. Lippincott Company, 1918. Pp. 209. $2.50.

This is the first of a series of monographs on Experimental Biology, under the editorship of Jacques Loeb, T. H. Morgan and W. J. V. Osterhout, which includes in its list of contributors many of the foremost students of biology in this country. "It is the aim of this monograph to show that the subject of animal conduct can be treated by the quantitative methods of the physicist, and that these methods lead to the forced movement or tropism theory of animal conduct, which was proposed by the writer thirty years ago, but which has only recently been carried to some degree of completion." The writer here summarizes the results of his own experimental work and that of his pupils on reactions to an electric current, to light, to the pull of gravity, to currents of water and air, to solid bodies, to chemicals, and to changes in temperature. The final chapters deal with instincts and memory images. There is a bibliography of 554 numbers.

ERNEST CARROLL MOORE. *Fifty Years of American Education.* Boston: Ginn and Company, 1917. Pp. 96. Eighty cents.

This little books was prepared for the firm of Ginn and Company as a kind of memorial of their fiftieth anniversary. In 1867 Edwin Ginn took desk room in a Boston office and started the enterprise which has developed into a national institution. But there is nothing commercial in this book. So far as we have been able to discover the name of the firm is not even mentioned in the body of the text. It is a compact, well-balanced survey of educational progress in this country in the past half-century, with especial emphasis on contemporary features of educational endeavor. As a clear and lucid summary of a most significant educational period, the volume is to be commended.

Journal of Educational Psychology
Vol. 9, 1918

Journal of Educational Psychology
Vol. 9, 1918.